PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/12, C07K 14 /705, A61K 38 /17, G01N 33 /68

A2

(11) International Publication Number: WO 00/05373

(43) International Publication Date: 3 February 2000 (03.02.00)

(21) International Application Number: PCT/US99/16484

(22) International Filing Date: 21 July 1999 (21.07.99)

(30) Priority Data:

60/093,630	21 July 1998 (21.07.98)	US
60/104,978	20 October 1998 (20.10.98)	US
09/245,041	5 February 1999 (05.02.99)	US

(71) Applicant: MILLENIUM PHARMACEUTICALS, INC. [US/US]; 640 Memorial Drive, Cambridge, MA 02139 (US).

(72) Inventors: MOORE, Karen; 34 Chandler Street, Maynard, MA 01754 (US). NAGLE, Deborah, L.; 370 Arlington Street, Watertown, MA 02172 (US).

(74) Agents: CORUZZI, Laura, A. et al.; Pennie & Edmonds LLP, 1155 Avenue of the Americas, New York, NY 10036 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY

(57) Abstract

The present invention relates to mammalian mahogany genes, including the human mahogany gene, which are novel genes involved in the control of mammalian body weight. The invention encompasses nucleotide sequences of the mahogany gene, host cell expression systems of the mahogany gene, and hosts which have been transformed by these expression systems, including transgenic animals. The invention also encompasses novel mahogany gene products, including mahogany proteins, polypeptides and peptides containing amino acid sequences mahogany proteins, fusion proteins of mahogany proteins polypeptides and peptides, and antibodies directed against such mahogany gene products. The present invention also relates to methods and compositions for the diagnosis and treatment of mammalian body weight disorders, including obesity, cachexia, and anorexia, and for the identification of subjects susceptible to such disorders. Further, the invention relates to methods of using the mahogany gene and gene products of the invention for the identification of compounds which modulate the expression of the mahogany gene and/or the activity of the mahogany gene product. Such compounds can be useful as therapeutic agents in the treatment of mammalian body weight disorders, including obesity, cachexia, and anorexia.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY

Priority of provisional application no. 60/093,630 filed on July 21, 1998 and of provisional application no. 60/104,978 filed on October 20, 1998, each of which is incorporated herein by reference in its entirety, is claimed under 35 U.S.C. § 119(e)(1).

1. INTRODUCTION

The present invention relates to mammalian mahogany genes, including the human mahogany gene, which are novel genes involved in the control of mammalian body weight. The invention encompasses nucleotide sequences of the mahogany gene, host cell expression systems of the mahogany gene, and hosts which have been transformed by these expression systems, including transgenic animals. The invention also encompasses novel mahogany gene products, including mahogany proteins, polypeptides and peptides containing amino acid sequences mahogany proteins, fusion proteins of mahogany proteins polypeptides and peptides, and antibodies directed against such mahogany gene products.

The present invention also relates to methods and compositions for the diagnosis and treatment of mammalian body weight disorders, including obesity, cachexia, and anorexia, and for the identification of subjects susceptible to such disorders. Further, the invention relates to methods of using the mahogany gene and gene products of the invention for the identification of compounds which modulate the expression of the mahogany gene and/or the activity of the mahogany gene product. Such compounds can be useful as therapeutic agents in the treatment of mammalian body weight disorders, including obesity, cachexia, and anorexia.

2. BACKGROUND OF THE INVENTION

Obesity represents the most prevalent of body weight disorders, and it is the most important nutritional disorder in the western world, with estimates of its prevalence ranging from 30% to 50% within the middle-aged population. Other body weight disorders, such as anorexia nervosa and bulimia nervosa, which together affect approximately 0.2% of the female population of the western world, also pose serious health threats. Further, such disorders as anorexia and cachexia (wasting) are also prominent features of other diseases such as cancer, cystic fibrosis, and AIDS.

Obesity, defined as an excess of body fat relative to lean body mass, also contributes to other diseases. For example, this disorder is responsible for increased incidence of diseases such as coronary artery disease, hypertension, stroke, diabetes, hyperlipidemia, and some cancers (See, *e.g.*, Nishina, P.M. *et al.*, 1994, Metab. 43: 554-558; Grundy, S.M. & Barnett, J.P., 1990, Dis. Mon. 36: 641-731). Obesity is not merely a behavioral problem, *i.e.*, the result of voluntary hyperphagia. Rather, the differential body composition observed between obese and normal subjects results from differences in both metabolism and neurologic/metabolic interactions. These differences seem to be, to some extent, due to differences in gene expression, and/or level of gene products or activity (Friedman, J.M. *et al.*, 1991, Mammalian Gene 1: 130-144).

The epidemiology of obesity strongly shows that the disorder exhibits inherited characteristics (Stunkard, 1990, N. Eng. J. Med. 322: 1438). Moll *et al.* have reported that, in many populations, obesity seems to be controlled by a few genetic loci (Moll *et al.*, 1991, Am. J. Hum. Gen. 49: 1243). In addition, human twin studies strongly suggest a substantial genetic basis in the control of body weight, with estimates of heritability of 80-90% (Simopoulos, A.P. &

Childs, B., eds., 1989, in "Genetic Variation and Nutrition in Obesity", World Review of Nutrition and Diabetes 63, S. Karger, Basel, Switzerland; Borjeson, M., 1976, Acta. Paediatr. Scand. 65: 279-287).

In other studies, non-obese persons who deliberately attempted to gain weight by systematically over-eating were found to be more resistant to such weight gain and able to maintain an elevated weight only by very high caloric intake. In contrast, spontaneously obese individuals are able to maintain their status with normal or only moderately elevated caloric intake. In addition, it is a commonplace experience in animal husbandry that different strains of swine, cattle, etc., have different predispositions to obesity. Studies of the genetics of human obesity, and of animal models of obesity demonstrate that obesity results from complex defective regulation of both food intake, food induced energy expenditure, and of the balance between lipid and lean body anabolism.

There are a number of genetic diseases in man and other species which feature obesity among their more prominent symptoms, along with, frequently, dysmorphic features and mental retardation. For example, Prader-Willi syndrome (PWS; reviewed in Knoll, J.H. *et al.*, 1993, Am. J. Med. Genet. 46: 2-6) affects approximately 1 in 20,000 live births, and involves poor neonatal muscle tone, facial and genital deformities, and generally obesity.

In addition to PWS, many other pleiotropic syndromes have been characterized which include obesity as a symptom. These syndromes are genetically straightforward, and appear to involve autosomal recessive alleles. Such diseases include, among others, Ahlstroem, Carpenter, Bardet-Biedl, Cohen, and Morgagni-Stewart-Monel Syndromes.

A number of models exists for the study of obesity (see, *e.g.*, Bray, G. A., 1992, Prog. Brain Res. 93: 333-341; and

Bray, G.A., 1989, Amer. J. Clin. Nutr. 5: 891-902). For example, animals having mutations which lead to syndromes that include obesity symptoms have also been identified. Attempts have been made to utilize such animals as models for the study of obesity, and the best studied animal models to date for genetic obesity are mice. For reviews, see, *e.g.*, Friedman, J.M. *et al.*, 1991, Mamm. Gen. 1: 130-144; Friedman, J.M. and Liebel, R.L., 1992, Cell 69: 217-220.

Studies utilizing mice have confirmed that obesity is a very complex trait with a high degree of heritability. Mutations at a number of loci have been identified which lead to obese phenotypes. These include the autosomal recessive mutations obese (*ob*), diabetes (*db*), fat (*fat*), and tubby (*tub*).

The dominant Yellow mutation (*Ay*) at the agouti locus causes a pleiotropic syndrome which causes moderate adult onset obesity, a yellow coat color, and a high incidence of tumor formation (Herberg, L. and Coleman, D.L., 1977, Metabolism 26:59), and an abnormal anatomic distribution of body fat (Coleman, D.L., 1978, Diabetologia 14:141-148). The mutation causes the widespread expression of a protein which is normally seen only in neonatal skin (Michaud, E. J. *et al.*, 1994, Genes Devel. 8:1463-1472). The agouti protein has been reported to be a competitive antagonist of α-MSH binding to the melanocortin receptors MC1-R and MC4-R *in vitro* (Lu *et al.*, 1996, Nature 371:799-802), and the authors speculated that de-regulated ubiquitous expression of agouti may lead to obesity by antagonism of melanocortin receptors expressed outside the hair follicles.

Mahogany (*mg*) and mahoganoid (*md*) are mutations that suppress the phenotypic effects of agouti protein *in vivo* (Lane and Green, 1960, J. Hered. 51: 228-230). The mahogany and mahoganoid mutation have been mapped to mouse chromosomes 2 and 16, respectively (Green, 1989, "Catalog of mutant genes

and polymorphic loci", pp. 12-403 in Genetic Variants and Strains of the Laboratory Mouse, Lyon, M. F. and Searle, A.G., eds., Oxford University Press, Oxford). Mutations of both *mg* and *md* have been shown to suppress the effects of agouti on obesity as well as on coat color (Miller *et al.*, 1997, Genetics 146: 1407-1415).

In summary, therefore, obesity, which poses a major, worldwide health problem, represents a complex, highly heritable trait. Given the severity, prevalence, and potential heterogeneity of such disorders, there exists a great need for the identification of those genes that participate in the control of body weight.

3. SUMMARY OF THE INVENTION

The present invention relates to the identification of novel nucleic acid molecules and proteins encoded by such nucleic acid molecules that are involved in the control of mammalian body weight, and which, further, are associated with mammalian body weight disorders such as obesity, cachexia, and anorexia. The nucleic acid molecules of the present invention represent the genes corresponding to the mammalian mahogany gene, including the human mahogany gene.

In particular, the compositions of the present invention include nucleic acid molecules which comprise the following sequences: (a) nucleotide sequences of the mahogany gene, including, e.g., murine mahogany sequences as shown in FIGS. 2A, 3B-D, 6A-B, 8A, and 9A, as well as allelic variants and homologs thereof, and human mahogany sequences, as shown, e.g., in FIGS. 10A, 18A, 19A and 20A, as well as allelic variants and homologs thereof; (b) nucleotide sequences that encode the mahogany gene product amino acid sequences, as shown, e.g.,in in FIGS. 2B, 8B, 9B, 10B, 17, 18B, 19B and 20B; (c) nucleotide sequences that encode portions of the mahogany gene product corresponding to its functional domains

and individual exons; (d) nucleotide sequences comprising the novel mahogany gene sequences disclosed herein that encode mutants of the mahogany gene product in which all or a part of one or more of the domains is deleted or altered, as shown, e.g., in FIG. 6; (e) nucleotide sequences that encode fusion proteins comprising the mahogany gene product, or one or more of its domains fused to a heterologous polypeptide; (f) nucleotide sequences within the mahogany gene, as well as chromosome sequences flanking the mahogany gene, see, e.g., FIG. 3, which can be utilized as part of the methods of the present invention for the diagnosis of mammalian body weight disorders, including obesity, cachexia, and anorexia, which are mediated by the mahogany gene, as well as for the identification of subjects susceptible to such disorders; (g) nucleic acid sequences that hybridize to the above described sequences under stringent or moderately stringent conditions, particularly human mg homologs. The nucleic acid molecules of the invention include, but are not limited to, cDNA and genomic DNA sequences of the mahogany gene.

The present invention also encompasses expression products of the nucleic acid molecules listed above; *i.e.*, proteins and/or polypeptides that are encoded by the above mahogany nucleic acid molecules.

Agonists and antagonists of the mahogany gene and/or gene product are also included in the present invention. Such agonists and antagonists will include, for example, small molecules, large molecules, and antibodies directed against the mahogany gene product. Agonists and antagonists of the invention also include nucleotide sequences, such as antisense and ribozyme molecules, and gene or regulatory sequence replacement constructs, that can be used to inhibit or enhance expression of the mahogany gene.

The present invention further encompasses cloning vectors, including expression vectors, that contain the

nucleic acid molecules of the invention and can be used to express those nucleic acid molecules in host organisms. The present invention also relates to host cells engineered to contain and/or express the nucleic acid molecules of the invention. Further, host organisms which have been transformed with these nucleic acid molecules are also encompassed in the present invention. Host organisms of the invention include organisms transformed with the cloning vectors described above, *e.g.*, transgenic animals, particularly non-human transgenic animals, and particularly transgenic non-human mammals.

The transgenic animals of the invention include animals that express a mutant variant or polymorphism of a mahogany gene, particularly a mutant variant or polymorphism of a mahogany gene that is associated with a weight disorder such as obesity, cachexia, or anorexia. The transgenic animals of the invention further include those that express a mahogany transgene at higher or lower levels than normal. The transgenic animals of the invention further include those which express the mahogany gene in all their cells, "mosaic" animals which express the mahogany gene in only some of their cells, and those in which the mahogany gene is selectively introduced into and expressed in a specific cell type(s). The transgenic animals of the invention also include "knock-out" animals. Knock-out animals comprise animals which have been engineered to no longer express the mahogany gene.

The present invention also relates to methods and compositions for the diagnosis of mammalian body weight disorders, including obesity, cachexia, and anorexia, as well as for the identification of subjects susceptible to such disorders. Such methods comprise, for example, measuring expression of the mahogany gene in a patient sample, or detecting a mutation in the mahogany gene in the genome of a mammal, including a human, suspected of exhibiting such a

weight disorder. The nucleic acid molecules of the invention can also be used as diagnostic hybridization probes, or as primers for diagnostic PCR analysis to identify of mahogany gene mutations, allelic variations, or regulatory defects, such as defects in the expression of the mahogany gene. Such diagnostic PCR analyses can be used to diagnose individuals with a body weight disorder associated with a particular mahogany gene mutation, allelic variation, or regulatory defect. Such diagnostic PCR analyses can also be used to identify individuals susceptible to such body weight disorders and hyperphagia.

Methods and compositions, including pharmaceutical compositions, for the treatment of body weight disorders such as obesity, cachexia, and anorexia are also included in the invention. Such methods and compositions are capable of modulating the level of mahogany gene expression and/or the level of activity of the mahogany gene product. Such methods include, for example, modulating the expression of the mahogany gene and/or the activity of the mahogany gene product for the treatment of a body weight disorder which is mediated by some other gene, for example by the agouti gene.

The invention still further relates to methods for identifying compounds which modulate the expression of the mammalian mahogany gene and/or the synthesis or activity of mammalian mahogany gene products. Such compounds include therapeutic compounds which can be used as pharmaceutical compositions to reduce or eliminate the symptoms of mammalian body weight disorders such as obesity, cachexia, and anorexia. Cellular and non-cellular assays are described that can be used to identify compounds that interact with the mahogany gene and/or gene product, *e.g.*, modulate the activity of the mahogany gene and/or bind to the mahogany gene product. Such cell-based assays of the invention

utilize cells, cell lines, or engineered cells or cell lines that express the mahogany gene product.

In one embodiment, such methods comprise contacting a compound to a cell that expresses a mahogany gene, measuring the level of mahogany gene expression, gene product expression, or gene product activity, and comparing this level to the level of mahogany gene expression, gene product expression, or gene product activity produced by the cell in the absence of the compound, such that if the level obtained in the presence of the compound differs from that obtained in its absence, a compound that modulates the expression of the mammalian mahogany gene and/or the synthesis or activity of mammalian mahogany gene products has been identified.

In an alternative embodiment, such methods comprise administering a compound to a host, *e.g.*, a transgenic animal that expresses a mahogany transgene or a mutant mahogany transgene, and measuring the level of mahogany gene expression, gene product expression, or gene product activity. The measured level is compared to the level of mahogany gene expression, gene product expression, or gene product activity in a host that is not exposed to the compound, such that if the level obtained when the host is exposed to the compound differs from that obtained when the host is not exposed to the compound, a compound that modulates the expression of the mammalian mahogany gene and/or the synthesis or activity of mammalian mahogany gene products, and/or the symptoms of a mammalian body weight disorder, such as obesity, cachexia, or anorexia, has been identified.

The Example presented in Section 6, below, describes the genetic and physical mapping of the mahogany gene to a specific 700 kb interval of mouse chromosome 2. The example presented in Section 7, below, describes the identification of a transcription unit within this chromosome interval,

referred to herein as the MG gene, which represents the mahogany gene. The expression and sequence analysis of this candidate mahogany gene is described in the example presented in Section 8, below. These experiments prove that the candidate gene MG is indeed the mahogany gene. The example presented in Section 9, below, presents data demonstrating that the mechanism of mahogany action is specific for diet-induced obestity, therefore supporting the use of mahogany antagonists as a specific therapeutic for treatment of diet-induced body weight disorders. The example presented in Section 10, below, presents the identification and characterization of the human mg gene, variants thereof and polypeptides encoded by the human mahogany sequences.

DEFINITIONS

As used herein, the following terms shall have the abbreviations indicated.

BAC, bacterial artificial chromosomes
bp, base pair(s)
EST, expressed sequence tag
mg, mahogany gene
RFLP, restriction fragment length polymorphism
RT-PCR, reverse transcriptase PCR
SSCP, single-stranded conformational polymorphism
SSLP, simple sequence length polymorphisms
STS, short tag sequence
YAC, yeast artificial chromosome

4. BRIEF DESCRIPTION OF THE FIGURES

FIG. 1. Physical map of the mahogany interval of mouse chromosome 2.

FIG. 2. Panel A(1)-A(3): cDNA nucleotide sequence of the wild-type (C57BL/6J) murine mahogany gene (SEQ ID NO: 1),

including the 5' and 3' untranslated regions, and Panel B: the derived amino acid sequence (SEQ ID NO: 2) of the mahogany gene product.

FIG. 3. Genomic structure and nucleotide sequences derived from the wild-type (C57BL/6J) mouse genomic regions containing the mg gene. Panel A, genomic structure; Panel B(1)-B(9), genomic sequence c56 (SEQ ID NO: 3); Panel C(1)-C(4), genomic sequence c96 (SEQ ID NO: 4); Panel D(1)-D(37), genomic sequence of c110/111 (SEQ ID NO: 5).

FIG. 4. Structural depiction of MG cDNA without introns. CUB=CUB domain, metal=metallothionin domain; T-transmembrane domain.

FIG. 5(1)-5(4). Nucleotide sequence of primers used to amplify each of the exons in the mg gene.

FIG. 6. Nucleotide sequence of the wild-type (SEQ ID NO: 6) and mahogany mutant (SEQ ID NO: 7) sequences in exon 15 of the MG gene. Bases shown in bold are deleted in Mg3J mutant mg.

FIG. 7. Differential 5' start sequences in the murine mahogany gene showing splice forms akm1003 and akm1004.

FIG. 8. Panel A, cDNA sequence (SEQ ID NO: 8) from one form of the differential 5' start site found in the murine (akm1003), Panel B, amino acid sequence (SEQ ID NO: 9) encoded by the cDNA of Panel A; Panel C, hydropathy plot of the akm1003 amino acid sequence.

FIG. 9. Panel A, cDNA sequence (SEQ ID NO: 10) from one form of the differential 5' start site found in the

murine (akm1004); Panel B, amino acid sequence (SEQ ID NO: 11) encoded by the cDNA of Panel A; Panel C, hydropathy plot of the akm1004 amino acid sequence.

FIG. 10. Nucleotide sequence (SEQ ID NO: 12) of a contig containing a portion of the human MG cDNA, panel A(1)-A(3) and the translated amino acid sequence (SEQ ID NO: 13), panel B.

FIG. 11. Effect of *mg* on *MC4r* -/- induced weight gain in females (FIG. 11A) and males (FIG. 11B); values depicted are the mean +/- SD within a designated time interval.

FIG. 12. Effect of *mg* on monogenic obese mutants $Lepr^{db}$ (FIG. 12A), *tub* (FIG. 12B), Cpe^{fat} (FIG. 12C), and on high fat diet induced obesity (FIG. 12D); the values indicated are the mean +/- SD of the weight length ratio for each animal.

FIG. 13. Genetic and physical map of the region surrounding the *mg* locus; all MIT markers are presented with shortened names, *e.g.*, D2MIT77 is indicated as D2M77; locations of loci which also mapped on the human cytogenetic map are indicated in parentheses after the gene symbol.

FIG. 13A. The genetic map of the *mg* gene region on the Millennium BSB mapping panel (Misumi, D.J. *et al.*, 1997, *Science 278*:135-138);

FIG. 13B. The genetic map obtained from crosses segregating *mg* mutant alleles;

FIG. 13C. The ~1 Mb BAC contig across the *mg* gene region of mouse Chromosome 2;

FIG. 13D. The transcriptional units identified in the *mg* region; the filled box indicates the *mg* gene,

whereas the hatched box is a member of the High Mobility Group (HMG) gene family which sits between coding exons 21 and 22 of the *mg* gene.

FIG. 14. Northern blot analysis with C3H/HeJ (lane 1), and three mutant alleles of *mg*: C3HeB/FeJ-*mg*^{3J} (Lane 2), LDJ/Le-*mg* (Lane 3), and C3H/HeJ-*mg*^{J} (Lane 4); the size marker is shown on the left, and hybridization with actin is shown below for loading comparisons.

FIG. 15. *In situ* hybridization data: FIG. 15A demonstrates widespread expression of *mg* throughout the mouse brain is seen in an antisense autoradiographic image of a C3H/HeJ brain at the level of the 3rd ventricle; decreased expression in *mg* mutants is documented in selected antisense darkfield images of 10 μm whole mount cross sections of the ventromedial hypothalamic nucleic (VMH) of C3H/HeJ (FIG. 15B), LDJ/Le-*mg* (FIG. 15C), and C3HeB/FeJ-*mg*^{3J} (FIG. 15D).

FIG. 16. Alignment of the MG protein sequence with its family members showing the transmembrane region (indicated in brackets) and cytoplasmic tail (FIG. 16A); and a schematic of the molecular modular architecture of MG (FIG 16B).

FIG. 17A-C. Sequence alignment of the predicted MG protein sequence (top) with the Attractin protein sequence. Characteristic MG domains are as indicated. See Section 10.2 for details.

FIG. 18A-B. Panel A: cDNA nucleotide sequence (SEQ ID NO: 14) of the long splice variant of the human ortholog of the mahogany gene, and Panel B: the derived amino acid sequence (SEQ ID NO: 15) of the mahogany gene product which it encodes.

FIG. 19A-B. Panel A: cDNA nucleotide sequence (SEQ ID NO: 16) of a shorter splice variant of the human ortholog of the mahogany gene, and Panel B: the derived amino acid sequence (SEQ ID NO: 17) of the mahogany gene product which it encodes.

FIG. 20A-B. Panel A: cDNA nucleotide sequence (SEQ ID NO: 18) of a second shorter splice variant of the human ortholog of the mahogany gene, and Panel B: the derived amino acid sequence (SEQ ID NO: 19) of the mahogany gene product which it encodes.

5. DETAILED DESCRIPTION OF THE INVENTION

Described herein is the identification of the novel mammalian mahogany (*mg*) gene, including the human mahogany gene, which is involved in the control of mammalian body weight. Also described are recombinant mammalian, including human mahogany DNA molecules, cloned genes, and degenerate variants thereof. The compositions of the present invention further include *mg* gene products (*e.g.*, proteins) that are encoded by the *mg* DNA molecules of the invention, and the modulation of *mg* gene expression and/or *mg* gene product activity in the treatment of mammalian body weight disorders, including obesity, cachexia, and anorexia. Also described herein are antibodies against *mg* gene products (*e.g.*, proteins), or conserved variants or fragments thereof, and nucleic acid probes useful for the identification of *mg* gene mutations, and the use of such nucleic acid probes in diagnosing mammalian body weight disorders, including obesity, cachexia, and anorexia. Further described are methods for the use of the *mg* gene and/or *mg* gene products in the identification of compounds which modulate the activity of the *mg* gene product.

5.1. THE MAHOGANY GENE

The mahogany genes are novel mammalian genes involved in the control of body weight. The nucleic acid sequences of the mahogany genes, including the murine mahogany gene sequences shown in FIGS. 2A, 3B-D, 6A-B, 8A, and 9A, as well as allelic variants and homologs thereof, and human mahogany sequences, as shown, e.g., in FIGS. 10A, 18A, 19A and 20A, as well as allelic variants and homologs thereof. The genomic sequence and structure, *i.e.*, the intron/exon structure, of the mahogany genes have also been elucidated, FIG. 3.

The mahogany gene nucleic acid molecules of the present invention comprise: (a) the DNA sequence shown in FIGS. 2A, 3, 6A-B, 8A, 9A, 10A, 18A, 19A or 20A, or any DNA sequence that encodes the amino acid sequence of the mahogany gene product shown in FIGS. 2B, 8B, 9B, 10B, 17, 18B, 19B or 20B; (b) nucleotide sequences comprising the novel mahogany sequences disclosed herein that encode mutants of the mahogany gene product in which all or a part of one or more of the domains is deleted or altered, as shown, e.g., FIG. 6; (c) nucleotide sequences that encode fusion proteins comprising a mahogany gene product, or one of its domains fused to a heterologous polypeptide; and (d) nucleotide sequences within a mahogany gene, nucleotide sequences on the chromosome flanking the mahogany gene, see, e.g., FIG. 3 and human genomic sequences syntenic to the sequences depicted in FIG. 3, which can be utilized as part of the methods of the invention for identifying and diagnosing individuals who exhibit or are susceptible to weight disorders, including obesity, cachexia, and anorexia.

The mahogany nucleotide sequences of the invention further comprise: (a) any nucleotide sequence that hybridizes to the complement of a nucleic acid molecule that encodes a mahogany gene product under highly stringent conditions , *e.g.*, hybridization to filter-bound DNA in 0.5 M

$NaHPO_4$, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1xSSC/0.1% SDS at 68°C (Ausubel F.M. *et al.*, eds., 1989, Current Protocols in Molecular Biology, Vol. I, Green Publishing Associates, Inc., and John Wiley & Sons, Inc., New York, at p. 2.10.3) particularly human mg sequences, FIG. 10; and (b) any nucleotide sequence that hybridizes to the complement of a nucleic acid molecule that encodes a mahogany gene product under less stringent conditions, such as moderately stringent conditions, *e.g.*,washing in 0.2xSSC/0.1% SDS at 42 °C (Ausubel *et al.*, 1989, *supra*), yet which still encodes a functionally equivalent mahogany gene product.

"Functionally equivalent", as utilized herein, refers to a gene product (*e.g.*, a protein) capable of exhibiting a substantially similar *in vivo* activity as the endogenous *mg* gene products encoded by the *mg* gene sequences described above. The *in vivo* activity of the *mg* gene product, as used herein, refers to the ability of the *mg* gene product, when present in an appropriate cell type, to ameliorate, prevent, or delay the appearance of the mahogany phenotype relative to its appearance when that cell type lacks a functional mahogany gene product.

The invention also includes nucleic acid molecules, preferably DNA molecules, that are the complements of the nucleotide sequences described above. Among the nucleic acid molecules of the invention are deoxyoligonucleotides ("oligos") which hybridize under highly stringent or moderately stringent conditions to the mahogany nucleic acid molecules described above. Exemplary highly stringent conditions may refer, *e.g.*, to washing in 6xSSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligos), 48°C (for 17-base oligos), 55°C (for 20-base oligos), and 60°C (for 23-base oligos). These nucleic acid molecules may encode or act as antisense molecules, useful, for example, in mahogany gene

regulation, and/or as antisense primers in amplification reactions of mahogany gene nucleic acid sequences. With respect to mahogany gene regulation, such techniques can be used to regulate, for example, weight disorders such as obesity, cachexia, or anorexia. Such sequences may also be used as part of ribozyme and/or triple helix sequences, which are also useful for mahogany gene regulation. Still further, such molecules may be used as components of diagnostic methods whereby, for example, the presence of a particular mahogany allele associated with a weight disorder, such as obesity, cachexia, or anorexia, may be detected. Among the molecules which can be used for diagnostic methods, such as those which involve amplification of genomic mahogany sequences, are primers or probes that can routinely be obtained using the genomic and cDNA sequences disclosed herein.

In one embodiment, the nucleic acid molecules of the invention do not include nucleic acid molecules that consist solely of the nucleotide sequence that encodes the attractin protein sequence depicted in FIG. 17A-C.

The mahogany nucleic acid sequences of the invention further include fragments of the nucleic acid sequences described above. For example, mahogany nucleic acid fragments can include fragments of at least 10, 12, 15, 20, 30, 40, 50, 100, 150, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000 or more nucleotides.

The nucleotide sequences of the present invention also include (a) DNA vectors that contain any of the foregoing mahogany coding sequences and/or their complements; (b) DNA expression vectors that contain any of the foregoing mahogany coding sequences operatively associated with a regulatory element that directs the expression of the coding sequences; and (c) genetically engineered host cells and organisms that

contain any of the foregoing mahogany coding sequences operatively associated with a regulatory element that directs the expression of the coding sequence in the host cell. As used herein, regulatory elements include, but are not limited to inducible and non-inducible promoters, enhancers, operators, and other elements known to those skilled in the art that drive and regulate gene expression. Such regulatory elements include, but are not limited to, the cytomegalovirus hCMV immediate early gene, the early or late promoters of SV40 adenovirus, the *lac* system, the *trp* system, the *TAC* system, the *TRC* system, the major operator and promoter regions of phage A, the control regions of fd coat protein, the promoter for 3'-phosphoglycerate kinase, the promoters of acid phosphatase, and the promoters of the yeast alpha-mating factors.

In addition to the mahogany gene sequences described above, homologs of such sequences, exhibiting extensive homology to one or more domains of the mahogany gene product can be present in other species. In a preferred embodiment, the mahogany gene homologue maps to a chromosomal region that is syntenic to the chromosomal region of the mahogany gene. In a particularly preferred embodiment, a human mahogany gene homologue sequence maps to a human chromosome region that is syntenic to the region of mouse chromosome 2 to which the murine mahogany gene maps, namely 20p15, and comprises the contiged human MG cDNA provided herein. Further, there can also exist homologue genes at other genetic loci within the genome of the same species which encode proteins having extensive homology to one or more domains of the mahogany gene product. Such mahogany homologs can include, for example, secreted forms of the mahogany sequences, see, e.g., Duke-Cohan, J.S. *et al.* (1998, *Proc. Natl. Acad. Sci. U.S.A.* *95*:11336-11341). Such sequences, can be used, for example, in the screening assays, described in Section 5.4.2 below,

for compounds that interact with the mahogany gene and/or its gene product and that may therefore be useful in treating and ameliorating body weight disorders.

Other mahogany homologs can be identified and readily isolated, without undue experimentation, by molecular biological techniques well known in the art, and are therefore within the scope of the present invention. As an example, in order to clone a human mahogany gene homologue using isolated murine mahogany gene sequences, such murine mahogany gene sequences may be labeled and used to screen a cDNA library constructed from mRNA obtained from appropriate cells or tissues derived from the organism (in this case, human) of interest. With respect to the cloning of such a human mahogany homologue, a human cDNA library may, for example be used for screening, such as a cDNA library obtained from mRNA isolated from brain tissues, particularly containing hypothalamic regions.

The hybridization washing conditions used should be of a lower stringency when the cDNA library is derived from an organism different from the type of organism from which the labeled sequence was derived. With respect to the cloning of a human mahogany homologue, for example, hybridization can be performed for 4 hours at 65°C using Amersham Rapid Hyb™ buffer (Cat. #RPN1639) according to manufacturer's protocol, followed by washing, with a final washing stringency of 1.0xSSC/0.1% SDS at 50°C for 20 minutes being preferred.

Low stringency conditions are well known to those of skill in the art, and will vary predictably depending on the specific organisms from which the library and the labeled sequences are derived. For guidance regarding such conditions see, for example, Sambrook *et al.*, 1989, Molecular Cloning, A Laboratory Manual, Cold Springs Harbor Press, N.Y.; and Ausubel *et al.*, 1989, Current Protocols in

Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y.

Alternatively, the labeled fragment may be used to screen a genomic library derived from the organism of interest, again, using appropriately stringent conditions.

Further, a mahogany gene homologue may be isolated from nucleic acid of the organism of interest by performing PCR using two degenerate oligonucleotide primer pools designed on the basis of amino acid sequences within the mahogany gene product disclosed herein. The template for the reaction may be cDNA obtained by reverse transcription of mRNA prepared from, for example, human or non-human cell lines or tissue known or suspected to express a mahogany gene allele.

The PCR product may be subcloned and sequenced to ensure that the amplified sequences represent the sequences of a mahogany gene nucleic acid sequence. The PCR fragment may then be used to isolate a full length cDNA clone by a variety of methods. For example, the amplified fragment may be labeled and used to screen a cDNA library, such as a bacteriophage cDNA library. Alternatively, the labeled fragment may be used to isolate genomic clones via the screening of a genomic library. This method has been used to isolate sequences encoding each of the murine MG gene exons as well as to isolate contigs containing the human MG sequences provided herein, FIG. 10.

PCR technology may also be utilized to isolate full length cDNA sequences. For example, RNA may be isolated, following standard procedures, from an appropriate cellular or tissue source (*i.e.*, one known, or suspected, to express the mahogany gene). A reverse transcription reaction may be performed on the RNA using an oligonucleotide primer specific for the most 5' end of the amplified fragment for the priming of the first strand synthesis. The resulting RNA/DNA hybrid may then be "tailed" with guanines using a standard terminal

transferase reaction, they hybrid may be digested with RNAase H, and second strand synthesis may then be primed with a poly-C primer. Thus, cDNA sequences upstream of the amplified fragment may easily be isolated. For a review of cloning strategies which may be used, see *e.g.*, Sambrook *et al.*, 1989 *supra*.

Mahogany gene sequences may additionally be used to isolate mutant mahogany alleles. Such mutant alleles may be isolated from individuals either known or proposed to have a phenotype which contributes to the symptoms of body weight disorders such as obesity, cachexia, or anorexia or disorders associated with hyperphagia. Mutant alleles and mutant allele products may then be utilized in the therapeutic and diagnostic systems described below. Additionally, such mahogany gene sequences can be used to detect mahogany gene regulatory (*e.g.* promoter) defects which can affect body weight.

A cDNA of a mutant mahogany gene may be isolated, for example, by using PCR, a technique which is well known to those of skill in the art. In this case, the first cDNA strand may be synthesized by hybridizing an oligo-dT oligonucleotide to mRNA isolated from tissue known or suspected to be expressed in an individual putatively carrying the mutant mahogany allele, and by extending the new strand with reverse transcriptase. The second strand of the cDNA is then synthesized using an oligonucleotide that hybridizes specifically to the 5' end of the normal gene. Using these two primers, the product is then amplified via PCR, cloned into a suitable vector, and subjected to DNA sequence analysis through methods well known to those of skill in the art. By comparing the DNA sequence of the mutant mahogany allele to that of the normal mahogany allele, the mutation(s) responsible for the loss of alteration of

activity of the mutant mahogany gene product can be ascertained.

Alternatively, a genomic library can be constructed using DNA obtained from an individual suspected of or known to carry the mutant mahogany allele, or a cDNA library can be constructed using RNA from a tissue known, or suspected to express the mutant mahogany allele. The normal mahogany gene or any suitable fragment thereof may then be labeled and used as a probe to identify the corresponding mutant mahogany allele is such libraries. Clones containing the mutant mahogany gene sequences may then be purified and subjected to sequence analysis according to methods well known to those of skill in the art.

Additionally, an expression library can be constructed utilizing cDNA synthesized from, for example, RNA isolated from a tissue known, or suspected to express a mutant mahogany allele in an individual suspected of or known to carry such a mutant allele. In this manner, gene products made by the putatively mutant tissue may be expressed and screened using standard antibody screening techniques in conjunction with antibodies raised against the normal mahogany gene product as described, below, in Section 5.3. For screening techniques, see, for example, Harlow, E. and Lane, eds., 1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor Press, Cold Spring Harbor. In cases where a mahogany mutation results in an expressed gene product with altered function (*e.g.*, as a result of a missense or a frameshift mutation) a polyclonal set of anti-mahogany gene product antibodies are likely to cross-react with the mutant mahogany gene product. Library clones detected via their reaction with such labeled antibodies can be purified and subjected to sequence analysis according to methods well known to those of skill in the art.

5.2. PROTEIN PRODUCTS OF THE MAHOGANY GENE

Mahogany gene products (e.g., proteins), polypeptides and peptide fragments, mutant, truncated, or deleted forms of the mahogany gene product, and/or fusion proteins of the mahogany gene product can be prepared for a variety of uses. For example, such gene products, or peptide fragments thereof, can be used for the generation of antibodies in diagnostic assays, or for the identification of other cellular or extracellular products involved in the regulation of mammalian body weight.

Mahogany gene products, also referred to herein as mahogany proteins, of the present invention include those gene products encoded by the mahogany gene sequences described in Section 5.1, above. For example, FIG. 2B, 8B and 9B depict murine mahogany amino acid sequences. Mahogany gene products also include human mahogany gene products as shown, e.g., in FIGS. 10B, 17B, 18B, 19B, and 20B.

In addition, mahogany gene products may include proteins that represent functionally equivalent gene products. Such an equivalent mahogany gene product may contain deletions, including internal deletions, additions, including additions yielding fusion proteins, or substitutions of amino acid residues within and/or adjacent to the amino acid sequence encoded by the mahogany gene sequences described, in Section 5.1, above, but that result in a "silent" change, in that the change produces a functionally equivalent mahogany gene product. Such amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and

glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid.

"Functionally equivalent", as utilized herein, refers to a gene product (*e.g.*, a protein) capable of exhibiting a substantially similar *in vivo* activity as the endogenous *mg* gene products encoded by the *mg* gene sequences described in Section 5.1, above. The *in vivo* activity of the *mg* gene product, as used herein, refers to the ability of the *mg* gene product, when present in an appropriate cell type, to ameliorate, prevent, or delay the appearance of the mahogany phenotype relative to its appearance when that cell type lacks a functional mahogany gene product.

Alternatively, where alteration of function is desired, deletion or non-conservative alterations can produce altered, including reduced-activity, mahogany gene products. Such alterations can, for example, alter one or more of the biological functions of the mahogany gene product. Further, such alterations can be selected so as to generate mahogany gene products that are better suited for expression, scale up, etc. in the host cells chosen. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

As another example, altered mahogany gene products can be engineered that correspond to mutants or variants of the mahogany gene product associated with mammalian weight disorders, such as obesity, cachexia, or anorexia. Altered mahogany gene products can also be engineered that correspond to mutants or variants of the mahogany gene product known to neutralize or ameliorate the symptoms of body weight disorders, such as obesity, cachexia, or anorexia, which are mediated by some other gene, including, but not limited to, body weight disorders mediated by the agouti gene.

Also within the scope of the present invention are peptides and/or proteins corresponding to one or more domains of the mahogany protein or any one of the individual exon encoded regions of the MG protein, as well as fusion proteins in which the full length mahogany protein, a mahogany peptide, or a truncated mahogany protein or peptide is fused to an unrelated heterologous protein. Such proteins and peptides can be designed on the basis of the mahogany nucleotide sequence disclosed in Section 5.1, above, and/or on the basis of the mahogany amino acid sequence disclosed in the Section.

The mahogany gene products of the invention further include fragments of the gene products described herein. For example, mahogany gene product fragments can include fragments of at least 10, 12, 15, 20, 30, 40, 50, 100, 150, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300 or more amino acids in length.

In one embodiment, it is understood that the gene products of the present invention do not include a gene product that consists solely of the amino acid sequence of the attractin polypeptide depicted in FIG. 17.

Fusion proteins of the invention include, but are not limited to, IgFc fusions which stabilize the mahogany protein or peptide and prolong half life *in vivo*; or fusions to any amino acid sequence that allows the fusion protein to be anchored to the cell membrane; or fusions to an enzyme, fluorescent protein, or luminescent protein which provides a marker function.

The mahogany gene products, peptide fragments thereof and fusion proteins thereof, may be produced by recombinant DNA technology using techniques well known in the art. Thus, methods for preparing the mahogany gene products, polypeptides, peptides, fusion peptide and fusion polypeptides of the invention by expressing nucleic acid

containing mahogany gene sequences are described herein. Methods that are well known to those skilled in the art can be used to construct expression vectors containing mahogany gene product coding sequences and appropriate transcriptional and translational control signals. These methods include, for example, *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. See, for example, the techniques described in Sambrook, *et al.*, 1989, *supra*, and Ausubel, *et al.*, 1989, *supra*. Alternatively, RNA capable of encoding mahogany gene product sequences may be chemically synthesized using, for example, synthesizers. See, for example, the techniques described in "Oligonucleotide Synthesis", 1984, Gait, ed., IRL Press, Oxford.

A variety of host-expression vector systems may be utilized to express the mahogany gene product coding sequences of the invention. Such host-expression systems represent vehicles by which the coding sequences of interest may be produced and subsequently purified, but also represent cells that may, when transformed or transfected with the appropriate nucleotide coding sequences, exhibit the mahogany gene product of the invention *in situ*. These include but are not limited to microorganisms such as bacteria (*e.g.*, *E. coli*, *B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing mahogany gene product coding sequences; yeast (*e.g.*, *Saccharomyces*, *Pichia*) transformed with recombinant yeast expression vectors containing the mahogany gene product coding sequences; insect cell systems infected with recombinant virus expression vectors (*e.g.*, baculovirus) containing the mahogany gene product coding sequences; plant cell systems infected with recombinant virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expres-

sion vectors (*e.g.*, Ti plasmid) containing mahogany gene product coding sequences; or mammalian cell systems (*e.g.*, COS, CHO, BHK, 293, 3T3) harboring recombinant expression constructs containing promoters derived from the genome of mammalian cells (*e.g.*, metallothionein promoter) or from mammalian viruses (*e.g.*, the adenovirus late promoter; the vaccinia virus 7.5K promoter).

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the mahogany gene product being expressed. For example, when a large quantity of such a protein is to be produced, for the generation of pharmaceutical compositions of mahogany gene product or for raising antibodies to mahogany gene product, for example, vectors that direct the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited, to the *E. coli* expression vector pUR278 (Ruther *et al.*, 1983, EMBO J. 2, 1791), in which the mahogany gene product coding sequence may be ligated individually into the vector in frame with the *lac Z* coding region so that a fusion protein is produced; pIN vectors (Inouye and Inouye, 1985, Nucleic Acids Res. 13, 3101-3109; Van Heeke and Schuster, 1989, J. Biol. Chem. 264, 5503-5509); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. The pGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned target gene product can be released from the GST moiety.

In an insect system, *Autographa californica*, nuclear polyhidrosis virus (AcNPV) is used as a vector to express

foreign genes. The virus grows in *Spodoptera frugiperda* cells. The mahogany gene product coding sequence may be cloned individually into non-essential regions (for example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter). Successful insertion of mahogany gene product coding sequence will result in inactivation of the polyhedrin gene and production of non-occluded recombinant virus (*i.e.*, virus lacking the proteinaceous coat coded for by the polyhedrin gene). These recombinant viruses are then used to infect *Spodoptera frugiperda* cells in which the inserted gene is expressed. (*e.g.*, see Smith, *et al.*, 1983, J. Virol. 46, 584; Smith, U.S. Patent No. 4,215,051).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the mahogany gene product coding sequence of interest may be ligated to an adenovirus transcription/translation control complex, *e.g.*, the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by *in vitro* or *in vivo* recombination. Insertion in a non-essential region of the viral genome (*e.g.*, region E1 or E3) will result in a recombinant virus that is viable and capable of expressing mahogany gene product in infected hosts. (*e.g.*, See Logan and Shenk, 1984, Proc. Natl. Acad. Sci. USA 81, 3655-3659). Specific initiation signals may also be required for efficient translation of inserted mahogany gene product coding sequences. These signals include the ATG initiation codon and adjacent sequences. In cases where an entire mahogany gene, including its own initiation codon and adjacent sequences, is inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only a portion of the mahogany gene coding sequence is inserted,

exogenous translational control signals, including, perhaps, the ATG initiation codon, must be provided. Furthermore, the initiation codon must be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements, transcription terminators, etc. (see Bittner, *et al.*, 1987, Methods in Enzymol. 153, 516-544).

In addition, a host cell strain may be chosen that modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (*e.g.*, glycosylation) and processing (*e.g.*, cleavage) of protein products may be important for the function of the protein. Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins and gene products. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells that possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. Such mammalian host cells include but are not limited to CHO, VERO, BHK, HeLa, COS, MDCK, 293, 3T3, and WI38.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines that stably express the mahogany gene product may be engineered. Rather than using expression vectors that contain viral origins of replication, host cells can be transformed with DNA controlled by appropriate expression control elements (*e.g.*, promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and

a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci that in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines that express the mahogany gene product. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that affect the endogenous activity of the mahogany gene product.

A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler, *et al.*, 1977, Cell 11, 223), hypoxanthine-guanine phosphoribosyltransferase (Szybalska and Szybalski, 1962, Proc. Natl. Acad. Sci. USA 48, 2026), and adenine phosphoribosyltransferase (Lowy, *et al.*, 1980, Cell 22, 817) genes can be employed in tk^-, $hgprt^-$ or $aprt^-$ cells, respectively. Also, antimetabolite resistance can be used as the basis of selection for the following genes: dhfr, which confers resistance to methotrexate (Wigler, *et al.*, 1980, Natl. Acad. Sci. USA 77, 3567; O'Hare, *et al.*, 1981, Proc. Natl. Acad. Sci. USA 78, 1527); gpt, which confers resistance to mycophenolic acid (Mulligan and Berg, 1981, Proc. Natl. Acad. Sci. USA 78, 2072); neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin, *et al.*, 1981, J. Mol. Biol. 150, 1); and hygro, which confers resistance to hygromycin (Santerre, *et al.*, 1984, Gene 30, 147).

Alternatively, the expression characteristic of an endogenous mahogany gene within a cell line or microorganism my be modified by inserting a heterologous DNA regulatory element into the genome of a stable cell line or cloned microorganism such that the inserted regulatory element is

operatively linked with the endogenous mahogany gene. For example, an endogenous mahogany gene which is normally "transcriptionally silent", *i.e.*, a mahogany gene which is normally not expressed, or is expressed only a very low levels in a cell line or microorganism, may be activated by inserting a regulatory element which is capable of promoting the expression of a normally expressed gene product in that cell line or microorganism. Alternatively, a transcriptionally silent, endogenous mahogany bene may be activated by insertion of a promiscuous regulatory element that works across cell types.

A heterologous regulatory element may be inserted into a stable cell line or cloned microorganism, such it is operatively linked with an endogenous mahogany gene, using techniques, such as targeted homologous recombination, which are well known to those of skill in the art, and described *e.g.*, in Chappel, U.S. Patent No. 4,215,051; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

Alternatively, any fusion protein may be readily purified by utilizing an antibody specific for the fusion protein being expressed. For example, a system described by Janknecht, *et al.* allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht, *et al.*, 1991, Proc. Natl. Acad. Sci. USA 88, 8972-8976). In this system, the gene of interest is subcloned into a vaccinia recombination plasmid such that the gene's open reading frame is translationally fused to an amino-terminal tag consisting of six histidine residues. Extracts from cells infected with recombinant vaccinia virus are loaded onto Ni^{2+}·nitriloacetic acid-agarose columns and

histidine-tagged proteins are selectively eluted with imidazole-containing buffers.

The mahogany gene products can also be expressed in transgenic animals. Animals of any species, including, but not limited to, mice, rats, rabbits, guinea pigs, pigs, micro-pigs, goats, sheep, and non-human primates, *e.g.*, baboons, monkeys, and chimpanzees may be used to generate mahogany transgenic animals. The term "transgenic," as used herein, refers to animals expressing mahogany gene sequences from a different species (*e.g.*, mice expressing human mahogany gene sequences), as well as animals that have been genetically engineered to over express endogenous (*i.e.*, same species) mahogany sequences or animals that have been genetically engineered to no longer express endogenous mahogany gene sequences (*i.e.*, "knock-out" animals), and their progeny.

Any technique known in the art may be used to introduce a mahogany gene transgene into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to pronuclear microinjection (Hoppe and Wagner, 1989, U.S. Pat. No. 4,873,191); retrovirus mediated gene transfer into germ lines (Van der Putten, *et al.*, 1985, Proc. Natl. Acad. Sci., USA 82, 6148-6152); gene targeting in embryonic stem cells (Thompson, *et al.*, 1989, Cell 56, 313-321); electroporation of embryos (Lo, 1983, Mol. Cell. Biol. 3, 1803-1814); and sperm-mediated gene transfer (Lavitrano *et al.*, 1989, Cell 57, 717-723) (For a review of such techniques, see Gordon, 1989, Transgenic Animals, Intl. Rev. Cytol. 115, 171-229)

Any technique known in the art may be used to produce transgenic animal clones containing a mahogany transgene, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal or adult cells induced to

quiescence (Campbell, *et al.*, 1996, Nature 380, 64-66; Wilmut, *et al.*, Nature 385, 810-813).

The present invention provides for transgenic animals that carry a mahogany transgene in all their cells, as well as animals that carry the transgene in some, but not all their cells, *i.e.*, mosaic animals. The transgene may be integrated as a single transgene or in concatamers, *e.g.*, head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, for example, the teaching of Lasko *et al.* (Lasko, *et al.*, 1992, Proc. Natl. Acad. Sci. USA 89, 6232-6236). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. When it is desired that the mahogany transgene be integrated into the chromosomal site of the endogenous mahogany gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous mahogany gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous mahogany gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous mahogany gene in only that cell type, by following, for example, the teaching of Gu, *et al.* (Gu, *et al.*, 1994, Science 265, 103-106). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

Once transgenic animals have been generated, the expression of the recombinant mahogany gene may be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to

analyze animal tissues to assay whether integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques that include but are not limited to Northern blot analysis of tissue samples obtained from the animal, *in situ* hybridization analysis, and RT-PCR (reverse transcriptase PCR). Samples of mahogany gene-expressing tissue, may also be evaluated immunocytochemically using antibodies specific for the mahogany transgene product.

5.3. ANTIBODIES TO MAHOGANY GENE PRODUCTS

Described herein are methods for the production of antibodies capable of specifically recognizing one or more *mg* gene product epitopes, or epitopes of conserved variants, or peptide fragments of the *mg* gene products. Further, antibodies that specifically recognize mutant forms of *mg* gene products, are encompassed by the invention.

Such antibodies may include, but are not limited to, polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, $F(ab')_2$ fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above. Such antibodies may be used, for example, in the detection of a *mg* gene product in an biological sample and may, therefore, be utilized as part of a diagnostic or prognostic technique whereby patients may be tested for abnormal levels of *mg* gene products, and/or for the presence of abnormal forms of such gene products. Such antibodies may also be utilized in conjunction with, for example, compound screening schemes, as described, below, in Section 5.4.2, for the evaluation of the effect of test compounds on *mg* gene product levels and/or activity. Additionally, such antibodies can be used in conjunction with the gene therapy techniques described,

below, in Section 5.4.3.2, to, for example, evaluate the normal and/or engineered mahogany-expressing cells prior to their introduction into the patient.

Anti-*mg* gene product antibodies may additionally be used in methods for inhibiting abnormal *mg* gene product activity. Thus, such antibodies may, therefore, be utilized as part of weight disorder treatment methods.

For the production of antibodies against a *mg* gene product, various host animals may be immunized by injection with a *mg* gene product, or a portion thereof. Such host animals may include, but are not limited to rabbits, mice, and rats, to name but a few. Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and *Corynebacterium parvum*.

Polyclonal antibodies are heterogeneous populations of antibody molecules derived from the sera of animals immunized with an antigen, such as a *mg* gene product, or an antigenic functional derivative thereof. For the production of polyclonal antibodies, host animals such as those described above, may be immunized by injection with *mg* gene product supplemented with adjuvants as also described above.

Monoclonal antibodies, which are homogeneous populations of antibodies to a particular antigen, may be obtained by any technique that provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique of Kohler and Milstein, (1975, Nature 256, 495-497; and U.S. Patent No. 4,376,110), the human B-cell hybridoma technique

(Kosbor *et al.*, 1983, Immunology Today 4, 72; Cole *et al.*, 1983, Proc. Natl. Acad. Sci. USA 80, 2026-2030), and the EBV-hybridoma technique (Cole *et al.*, 1985, Monoclonal Antibodies And Cancer Therapy, Alan R. Liss, Inc., pp. 77-96). Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the mAb of this invention may be cultivated *in vitro* or *in vivo*. Production of high titers of mAbs *in vivo* makes this the presently preferred method of production.

In addition, techniques developed for the production of "chimeric antibodies" (Morrison, *et al.*, 1984, Proc. Natl. Acad. Sci., 81, 6851-6855; Neuberger, *et al.*, 1984, Nature 312, 604-608; Takeda, *et al.*, 1985, Nature, 314, 452-454) by splicing the genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region. (See, *e.g.*, Cabilly et al., U.S. Patent No. 4,816,567; and Boss et al., U.S. Patent No. 4,816397, which are incorporated herein by reference in their entirety.)

In addition, techniques have been developed for the production of humanized antibodies. (See, *e.g.*, Queen, U.S. Patent No. 5,585,089, which is incorporated herein by reference in its entirety.) An immunoglobulin light or heavy chain variable region consists of a "framework" region interrupted by three hypervariable regions, referred to as complementarily determining regions (CDRs). The extent of the framework region and CDRs have been precisely defined (see, "Sequences of Proteins of Immunological Interest", Kabat, E. et al., U.S. Department of Health and Human Services (1983). Briefly, humanized antibodies are antibody

molecules from non-human species having one or more CDRs from the non-human species and a framework region from a human immunoglobulin molecule.

Alternatively, techniques described for the production of single chain antibodies (U.S. Patent 4,946,778; Bird, 1988, Science 242, 423-426; Huston, *et al.*, 1988, Proc. Natl. Acad. Sci. USA 85, 5879-5883; and Ward, *et al.*, 1989, Nature 334, 544-546) can be adapted to produce single chain antibodies against mahogany gene products. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide.

Antibody fragments that recognize specific epitopes may be generated by known techniques. For example, such fragments include but are not limited to: the $F(ab')_2$ fragments, which can be produced by pepsin digestion of the antibody molecule and the Fab fragments, which can be generated by reducing the disulfide bridges of the $F(ab')_2$ fragments. Alternatively, Fab expression libraries may be constructed (Huse, *et al.*, 1989, Science, 246, 1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity.

5.4. USES OF THE MAHOGANY GENES, GENE PRODUCTS, AND ANTIBODIES

Described herein are various applications of the mahogany genes, of the mahogany gene products, including peptide fragments thereof, and of antibodies directed against mahogany gene products and peptide fragments thereof. Such applications include, for example, prognostic and diagnostic evaluation of body weight disorders and the identification of subjects with a predisposition to such disorders, as described below, in Section 5.4.1. Additionally, such applications include methods for the treatment of body weight

and body weight disorders, as described, below, in Section 5.4.2, and for the identification of compounds which modulate the expression of the mahogany gene and/or the activity of the mahogany gene product, as described in Section 5.4.3, below. Such compounds can include, for example, other cellular products which are involved in body weight regulation. These compounds can be used, for example, in the amelioration of body weight disorders, including obesity, cachexia, and anorexia.

While, for clarity, the uses described in this section are primarily uses related to body weight disorder abnormalities, it is to be noted that each of the diagnostic and therapeutic treatments described herein can additionally be utilized in connection with other defects associated with the mahogany gene, such as hyperpigmentation, hyperphagia and other disorders resulting in increased metabolic rates.

5.4.1. DIAGNOSIS OF BODY WEIGHT DISORDER ABNORMALITIES

A variety of methods can be employed for the diagnostic and prognostic evaluation of body weight disorders, including obesity, cachexia, and anorexia, and for the identification of subjects having a predisposition to such disorders.

Such methods may, for example, utilize reagents such as the mahogany gene nucleotide sequences described in Section 5.1, and antibodies directed against mahogany gene products, including peptide fragments thereof, as described, above, in Section 5.3. Specifically, such reagents may be used, for example, for:

(1) the detection of the presence of mahogany gene mutations, or the detection of either over- or under-expression of mahogany gene relative to levels of mahogany expression in a wild-type, non-body weight disorder state

which correlates with certain body weight disorders or susceptibility toward such body weight disorders;

(2) the detection of over- or under-abundance of mahogany gene product relative to the abundance of mahogany gene product in a wild-type non-body weight disorder state which correlates with certain body weight disorders or susceptibility toward such body weight disorders; and

(3) the detection of an aberrant level of mahogany gene product activity relative to mahogany gene product activity levels in a wild-type, non-body weight disorder state which correlates with certain body weight disorders or susceptibility toward such body weight disorders.

Mahogany gene nucleotide sequences can, for example, be used to diagnose a body weight disorder using, for example, the techniques for detecting mutations in the mahogany gene described above in Section 5.1, above.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one specific mahogany gene nucleic acid or anti-mahogany gene product antibody reagent described herein, which may be conveniently used, *e.g.*, in clinical settings, to screen and diagnose patients exhibiting body weight disorder abnormalities, and to screen those individuals exhibiting a predisposition to developing a body weight disorder abnormality.

For the detection of mahogany gene mutations, any nucleated cell can be used as a starting source for genomic nucleic acid. For the detection of mahogany gene expression or mahogany gene products, any cell type or tissue in which the mahogany gene is expressed may be utilized, such as, for example, tissues or cells shown herein to express the MG gene.

Nucleic acid-based detection techniques are described, below, in Section 5.4.1.1. Peptide detection techniques are described, below, in Section 5.4.1.2.

5.4.1.1. DETECTION OF MAHOGANY GENE NUCLEIC ACID MOLECULES

Mutations or polymorphisms within the mahogany gene can be detected by utilizing a number of techniques. Nucleic acid from any nucleated cell can be used as the starting point for such assay techniques, and may be isolated according to standard nucleic acid preparation procedures which are well known to those of skill in the art.

Genomic DNA may be used in hybridization or amplification assays of biological samples to detect abnormalities involving mahogany gene structure, including point mutations, insertions, deletions and chromosomal rearrangements. Such assays may include, but are not limited to, Southern analyses, single stranded conformation polymorphism analyses (SSCP), and PCR analyses.

Diagnostic methods for the detection of mahogany gene-specific mutations can involve for example, contacting and incubating nucleic acids obtained from a sample, *e.g.*, derived from a patient sample or other appropriate cellular source with one or more labeled nucleic acid reagents including recombinant DNA molecules, cloned genes or degenerate variants thereof, such as described in Section 5.1, above, under conditions favorable for the specific annealing of these reagents to their complementary sequences within or flanking the mahogany gene. Preferably, the lengths of these nucleic acid reagents are at least 15 to 30 nucleotides.

After incubation, all non-annealed nucleic acids are removed from the nucleic acid:mahogany molecule hybrid. The presence of nucleic acids that have hybridized, if any such

molecules exist, is then detected. Using such a detection scheme, the nucleic acid from the cell type or tissue of interest can be immobilized, for example, to a solid support such as a membrane, or a plastic surface such as that on a microtiter plate or polystyrene beads. In this case, after incubation, non-annealed, labeled nucleic acid reagents of the type described in Section 5.1 are easily removed. Detection of the remaining, annealed, labeled mahogany nucleic acid reagents is accomplished using standard techniques well-known to those in the art. The mahogany gene sequences to which the nucleic acid reagents have annealed can be compared to the annealing pattern expected from a normal mahogany gene sequence in order to determine whether a mahogany gene mutation is present.

In a preferred embodiment, mahogany gene mutations or polymorphisms can be detected by using a microassay of mahogany nucleic acid sequences immobilized to a substrate or "gene chip" (see, *e.g.* Cronin, et al., 1996, Human Mutation 7:244-255).

Alternative diagnostic methods for the detection of mahogany gene specific nucleic acid molecules, in patient samples or other appropriate cell sources, may involve their amplification, *e.g.*, by PCR (the experimental embodiment set forth in Mullis, 1987, U.S. Patent No. 4,683,202), followed by the analysis of the amplified molecules using techniques well known to those of skill in the art, such as, for example, those listed above. The resulting amplified sequences can be compared to those that would be expected if the nucleic acid being amplified contained only normal copies of the mahogany gene in order to determine whether a mahogany gene mutation exists.

Among those mahogany nucleic acid sequences which are preferred for such amplification-related diagnostic screening analyses are oligonucleotide primers which amplify mahogany

exon sequences. The sequences of such oligonucleotide primers are, therefore, preferably derived from mahogany intron sequences so that the entire exon, or coding region, can be analyzed as discussed below. Primer pairs useful for amplification of mahogany exons are preferably derived from adjacent introns. Appropriate primer pairs can be chosen such that each of the 25 mahogany exons are amplified. Primers for the amplification of mahogany exons can be routinely designed by one of ordinary skill in the art by utilizing the exon and intron sequences of mahogany shown in Figures, particularly FIGS. 3 and 5.

Additional mahogany nucleic acid sequences which are preferred for such amplification-related analyses are those which will detect the presence of a mahogany polymorphism which doffers from the consensus mahogany sequence depicted in Figures, particularly those that detect the polymorphism identified in exon 15 (Figure 7). Such polymorphisms include ones which represent mutations associated with body weight disorders such as obesity, cachexia, or anorexia.

Further, well-known genotyping techniques can be performed to type polymorphisms that are in close proximity to mutations in the mahogany gene itself, including mutations associated with weight disorders such as obesity, cachexia, or anorexia. Such polymorphisms can be used to identify individuals in families likely to carry mutations in the mahogany gene. If a polymorphism exhibits linkage disequilibrium with mutations in the mahogany gene, the polymorphism can also be used to identify individuals in the general population who are likely to carry such mutations. Polymorphisms that can be used in this way include restriction fragment length polymorphisms (RFLPs), which involve sequence variations in restriction enzyme target sequences, single-base polymorphisms, and simple sequence length polymorphisms (SSLPs).

For example, Weber (U.S. Pat. No. 5,075,217) describes a DNA marker based on length polymorphisms in blocks of (dC-dA)n-(dG-dT)n short tandem repeats. The average separation of (dC-dA)n-(dG-dT)n blocks is estimated to be 30,000-60,000 bp. Markers that are so closely spaced exhibit a high frequency co-inheritance, and are extremely useful in the identification of genetic mutations, such as, for example, mutations within the mahogany gene, and the diagnosis of diseases and disorders related to mutations in the mahogany gene.

Also, Caskey *et al.* (U.S. Pat.No. 5,364,759) describe a DNA profiling assay for detecting short tri and tetra nucleotide repeat sequences. The process includes extracting the DNA of interest, such as the mahogany gene, amplifying the extracted DNA, and labelling the repeat sequences to form a genotypic map of the individual's DNA.

A mahogany probe could additionally be used to directly identify RFLPs. Further, a mahogany probe or primers derived from the mahogany sequence could be used to isolate genomic clones such as YACs, BACs, PACs, cosmids, phage, or plasmids. The DNA contained in these clones can be screened for single-base polymorphisms or SSLPs using standard hybridization or sequencing procedures.

The level of mahogany gene expression can also be assayed. For example, RNA from a cell type or tissue known, or suspected, to express the mahogany gene, such as muscle, brain, kidney, testes, heart, liver, lung, skin, hypothalamus, spleen, and adipose tissue may be isolated and tested utilizing hybridization or PCR techniques such as are described, above. The isolated cells can be derived from cell culture or from a patient. The analysis of cells taken from culture may be a necessary step in the assessment of cells to be used as part of a cell-based gene therapy technique or, alternatively, to test the effect of compounds

on the expression of the mahogany gene. Such analyses may reveal both quantitative and qualitative aspects of the expression pattern of the mahogany gene, including activation or inactivation of mahogany gene expression.

In one embodiment of such a detection scheme, a cDNA molecule is synthesized from an RNA molecule of interest (*e.g.*, by reverse transcription of the RNA molecule into cDNA). All or part of the resulting cDNA is then used as the template for a nucleic acid amplification reaction, such as a PCR amplification reaction, or the like. The nucleic acid reagents used as synthesis initiation reagents (*e.g.*, primers) in the reverse transcription and nucleic acid amplification steps of this method are chosen from among the mahogany gene nucleic acid reagents described in Section 5.1. The preferred lengths of such nucleic acid reagents are at least 9-30 nucleotides.

For detection of the amplified product, the nucleic acid amplification may be performed using radioactively or non-radioactively labeled nucleotides. Alternatively, enough amplified product may be made such that the product may be visualized by standard ethidium bromide staining or by utilizing any other suitable nucleic acid staining method.

As an alternative to amplification techniques, standard Northern analyses can be performed to determine the level of mRNA expression of the mahogany gene, if a sufficient quantity of the appropriate cells can be obtained.

Additionally, it is possible to perform such mahogany gene expression assays "in situ", *i.e.*, directly upon tissue sections (fixed and/or frozen) of patient tissue obtained from biopsies or resections, such that no nucleic acid purification is necessary. Nucleic acid reagents such as those described in Section 5.1 may be used as probes and/or primers for such in situ procedures (see, for example, Nuovo,

G.J., 1992, "PCR In Situ Hybridization: Protocols And Applications", Raven Press, NY).

5

5.4.1.2. DETECTION OF MAHOGANY GENE PRODUCTS

Mahogany gene products, including both wild-type and mutant mahogany gene products, conserved variants, and polypeptide fragments thereof, which are discussed, above, in Section 5.2, may be detected using antibodies which are directed against such mahogany gene products. Such antibodies, which are discussed in Section 5.3, below, may thereby be used as diagnostics and prognostics for a body weight disorder. Such methods may be used to detect abnormalities in the level of mahogany gene expression or of mahogany gene product synthesis, or abnormalities in the structure, temporal expression, and/or physical location of mahogany gene product. The antibodies and immunoassay methods described herein have, for example, important *in vitro* applications in assessing the efficacy of treatments for body weight disorders such as obesity, cachexia, and anorexia. Antibodies, or fragments of antibodies, such as those described below, may be used to screen potentially therapeutic compounds *in vitro* to determine their effects on mahogany gene expression and mahogany gene product production. The compounds that have beneficial effects on body weight disorders, such as obesity, cachexia, and anorexia, can thereby be identified, and a therapeutically effective dose determined.

In vitro immunoassays may also be used, for example, to assess the efficacy of cell-based gene therapy for a body weight disorders, including obesity, cachexia, and anorexia. Antibodies directed against mahogany gene products may be used *in vitro* to determine, for example, the level of

mahogany gene expression achieved in cells genetically engineered to produce mahogany gene product. In the case of intracellular mahogany gene products, such an assessment is done, preferably, using cell lysates or extracts. Such analysis will allow for a determination of the number of transformed cells necessary to achieve therapeutic efficacy *in vivo*, as well as optimization of the gene replacement protocol.

The tissue or cell type to be analyzed will generally include those that are known, or suspected, to express the mahogany gene. The protein isolation methods employed herein may, for example, be such as those described in Harlow and Lane (1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York). The isolated cells can be derived from cell culture or from a patient. The analysis of cells taken from culture may be a necessary step in the assessment of cells to be used as part of a cell-based gene therapy technique or, alternatively, to test the effect of compounds on the expression of the mahogany gene.

Preferred diagnostic methods for the detection of mahogany gene products, conserved variants or peptide fragments thereof, may involve, for example, immunoassays wherein the mahogany gene products or conserved variants or peptide fragments are detected by their interaction with an anti-mahogany gene product-specific antibody.

For example, antibodies, or fragments of antibodies, such as those described, above, in Section 5.3, may be used to quantitatively or qualitatively detect the presence of mahogany gene products or conserved variants or peptide fragments thereof. This can be accomplished, for example, by immunofluorescence techniques employing a fluorescently labeled antibody (see below, this Section) coupled with light microscopic, flow cytometric, or fluorimetric detection.

Such techniques are especially preferred for mahogany gene products that are expressed on the cell surface.

The antibodies (or fragments thereof) useful in the present invention may, additionally, be employed histologically, as in immunofluorescence or immunoelectron microscopy, for *in situ* detection of mahogany gene products, conserved variants or peptide fragments thereof. *In situ* detection may be accomplished by removing a histological specimen from a patient, and applying thereto a labeled antibody that binds to a mahogany polypeptide. The antibody (or fragment) is preferably applied by overlaying the labeled antibody (or fragment) onto a biological sample. Through the use of such a procedure, it is possible to determine not only the presence of the mahogany gene product, conserved variants or peptide fragments, but also its distribution in the examined tissue. Using the present invention, those of ordinary skill will readily recognize that any of a wide variety of histological methods (such as staining procedures) can be modified in order to achieve *in situ* detection of a mahogany gene product.

Immunoassays for mahogany gene products, conserved variants, or peptide fragments thereof will typically comprise: (1) incubating a sample, such as a biological fluid, a tissue extract, freshly harvested cells, or lysates of cells in the presence of a detectably labeled antibody capable of identifying mahogany gene products, conserved variants or peptide fragments thereof; and (2) detecting the bound antibody by any of a number of techniques well-known in the art.

The biological sample may be brought in contact with and immobilized onto a solid phase support or carrier, such as nitrocellulose, that is capable of immobilizing cells, cell particles or soluble proteins. The support may then be washed with suitable buffers followed by treatment with the

detectably labeled mahogany gene product specific antibody. The solid phase support may then be washed with the buffer a second time to remove unbound antibody. The amount of bound label on the solid support may then be detected by conventional means.

By "solid phase support or carrier" is intended any support capable of binding an antigen or an antibody. Well-known supports or carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, gabbros, and magnetite. The nature of the carrier can be either soluble to some extent or insoluble for the purposes of the present invention. The support material may have virtually any possible structural configuration so long as the coupled molecule is capable of binding to an antigen or antibody. Thus, the support configuration may be spherical, as in a bead, or cylindrical, as in the inside surface of a test tube, or the external surface of a rod. Alternatively, the surface may be flat such as a sheet, test strip, etc. Preferred supports include polystyrene beads. Those skilled in the art will know many other suitable carriers for binding antibody or antigen, or will be able to ascertain the same by use of routine experimentation.

One of the ways in which the mahogany gene product-specific antibody can be detectably labeled is by linking the same to an enzyme, such as for use in an enzyme immunoassay (EIA) (Voller, A., "The Enzyme Linked Immunosorbent Assay (ELISA)", 1978, Diagnostic Horizons 2, 1-7, Microbiological Associates Quarterly Publication, Walkersville, MD); Voller, A. *et al.*, 1978, J. Clin. Pathol. 31, 507-520; Butler, J.E., 1981, Meth. Enzymol. 73, 482-523; Maggio, E. (ed.), 1980, Enzyme Immunoassay, CRC Press, Boca Raton, FL,; Ishikawa, E. *et al.*, (eds.), 1981, Enzyme Immunoassay, Kgaku Shoin, Tokyo). The enzyme which is bound to the antibody will react

with an appropriate substrate, preferably a chromogenic substrate, in such a manner as to produce a chemical moiety that can be detected, for example, by spectrophotometric, fluorimetric or by visual means. Enzymes that can be used to detectably label the antibody include, but are not limited to, malate dehydrogenase, staphylococcal nuclease, delta-5-steroid isomerase, yeast alcohol dehydrogenase, α-glycerophosphate, dehydrogenase, triose phosphate isomerase, horseradish peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, β-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase and acetylcholinesterase. The detection can be accomplished by colorimetric methods that employ a chromogenic substrate for the enzyme. Detection may also be accomplished by visual comparison of the extent of enzymatic reaction of a substrate in comparison with similarly prepared standards.

Detection may also be accomplished using any of a variety of other immunoassays. For example, by radioactively labeling the antibodies or antibody fragments, it is possible to detect mahogany gene products through the use of a radioimmunoassay (RIA) (see, for example, Weintraub, B., Principles of Radioimmunoassays, Seventh Training Course on Radioligand Assay Techniques, The Endocrine Society, March, 1986). The radioactive isotope can be detected by such means as the use of a gamma counter or a scintillation counter or by autoradiography.

It is also possible to label the antibody with a fluorescent compound. When the fluorescently labeled antibody is exposed to light of the proper wave length, its presence can then be detected due to fluorescence. Among the most commonly used fluorescent labeling compounds are fluorescein isothiocyanate, rhodamine, phycoerythrin, phycocyanin, allophycocyanin, *o*-phthaldehyde and fluorescamine.

The antibody can also be detectably labeled using fluorescence emitting metals such as ^{152}Eu, or others of the lanthanide series. These metals can be attached to the antibody using such metal chelating groups as diethylenetriaminepentacetic acid (DTPA) or ethylenediaminetetraacetic acid (EDTA).

The antibody also can be detectably labeled by coupling it to a chemiluminescent compound. The presence of the chemiluminescent-tagged antibody is then determined by detecting the presence of luminescence that arises during the course of a chemical reaction. Examples of particularly useful chemiluminescent labeling compounds are luminol, isoluminol, theromatic acridinium ester, imidazole, acridinium salt and oxalate ester.

Likewise, a bioluminescent compound may be used to label the antibody of the present invention. Bioluminescence is a type of chemiluminescence found in biological systems in which a catalytic protein increases the efficiency of the chemiluminescent reaction. The presence of a bioluminescent protein is determined by detecting the presence of luminescence. Important bioluminescent compounds for purposes of labeling are luciferin, luciferase and aequorin.

5.4.2. SCREENING ASSAYS FOR COMPOUNDS THAT INTERACT WITH THE MAHOGANY GENE OR GENE PRODUCT

The following assays are designed to identify compounds that bind to a mahogany gene product, compounds that bind to proteins, or portions of proteins that interact with a mahogany gene product, compounds that interfere with the interaction of a mahogany gene product with proteins and compounds that modulate the activity of the mahogany gene (*i.e.*, modulate the level of mahogany gene expression and/or modulate the level of mahogany gene product activity). Assays may additionally be utilized that identify compounds

that bind to mahogany gene regulatory sequences (*e.g.*, promoter sequences; see *e.g.*, Platt, 1994, J. Biol. Chem. 269, 28558-28562), which is incorporated herein by reference in its entirety, and that can modulate the level of mahogany gene expression. Such compounds may include, but are not limited to, small organic molecules, such as ones that are able to cross the blood-brain barrier, gain to and/or entry into an appropriate cell and affect expression of the mahogany gene or some other gene involved in the body weight regulatory pathway, or intracellular proteins.

Methods for the identification of such proteins are described, below, in Section 5.4.2.2. Such proteins may be involved in the control and/or regulation of body weight. Further, among these compounds are compounds that affect the level of mahogany gene expression and/or mahogany gene product activity and that can be used in the therapeutic treatment of body weight disorders, including obesity, cachexia, and anorexia, as described, below, in Section 5.9.

Compounds may include, but are not limited to, peptides such as, for example, soluble peptides, including but not limited to, Ig-tailed fusion peptides, and members of random peptide libraries; (see, *e.g.*, Lam, *et al.*, 1991, Nature 354, 82-84; Houghten, *et al.*, 1991, Nature 354, 84-86), and combinatorial chemistry-derived molecular library made of D- and/or L- configuration amino acids, phosphopeptides (including, but not limited to members of random or partially degenerate, directed phosphopeptide libraries; see, *e.g.*, Songyang, *et al.*, 1993, Cell 72, 767-778), antibodies (including, but not limited to, polyclonal, monoclonal, humanized, anti-idiotypic, chimeric or single chain antibodies, and FAb, $F(ab')_2$ and FAb expression library fragments, and epitope-binding fragments thereof), and small organic or inorganic molecules.

Compounds identified via assays such as those described herein may be useful, for example, in elaborating the biological function of the mahogany gene product and for ameliorating body weight disorders, such as obesity, cachexia, or anorexia. Assays for testing the effectiveness of compounds identified by, for example, techniques such as those described in Sections 5.4.2.1-5.4.2.3, are discussed, below, in Section 5.4.2.4.

5.4.2.1. *IN VITRO* SCREENING ASSAYS FOR COMPOUNDS THAT BIND TO THE MAHOGANY GENE PRODUCT

In vitro systems may be designed to identify compounds capable of binding the mahogany gene products of the invention. Compounds identified may be useful, for example, in modulating the activity of unimpaired and/or mutant mahogany gene products, may be useful in elaborating the biological function of the mahogany gene product, may be utilized in screens for identifying compounds that disrupt normal mahogany gene product interactions, or may in themselves disrupt such interactions.

The principle of the assays used to identify compounds that bind to the mahogany gene product involves preparing a reaction mixture of the mahogany gene product and the test compound under conditions and for a time sufficient to allow the two components to interact and bind, thus forming a complex that can be removed and/or detected in the reaction mixture. These assays can be conducted in a variety of ways. For example, one method to conduct such an assay involves anchoring a mahogany gene product or a test substance onto a solid support and detecting mahogany gene product/test compound complexes formed on the solid support at the end of the reaction. In one embodiment of such a method, the mahogany gene product may be anchored onto a solid support,

and the test compound, which is not anchored, may be labeled, either directly or indirectly.

In practice, microtiter plates are conveniently utilized as the solid support. The anchored component may be immobilized by non-covalent or covalent attachments. Non-covalent attachment may be accomplished by simply coating the solid surface with a solution of the protein and drying. Alternatively, an immobilized antibody, preferably a monoclonal antibody, specific for the protein to be immobilized may be used to anchor the protein to the solid surface. The surfaces may be prepared in advance and stored.

In order to conduct the assay, the non-immobilized component is added to the coated surface containing the anchored component. After the reaction is complete, unreacted components are removed (*e.g.*, by washing) under conditions such that any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of ways. Where the previously non-immobilized component is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the previously non-immobilized component is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface; *e.g.*, using a labeled antibody specific for the previously non-immobilized component (the antibody, in turn, may be directly labeled or indirectly labeled with a labeled anti-Ig antibody).

Alternatively, a reaction can be conducted in a liquid phase, the reaction products separated from unreacted components, and complexes detected; *e.g.*, using an immobilized antibody specific for mahogany gene product or the test compound to anchor any complexes formed in solution, and a labeled antibody specific for the other component of the possible complex to detect anchored complexes.

5.4.2.2. ASSAYS FOR PROTEINS THAT INTERACT WITH THE MAHOGANY GENE PRODUCT

Any method suitable for detecting protein-protein interactions may be employed for identifying mahogany gene product-protein interactions.

Among the traditional methods that may be employed are co-immunoprecipitation, cross-linking and co-purification through gradients or chromatographic columns. Utilizing procedures such as these allows for the identification of proteins that interact with mahogany gene products. Such proteins can include, but are not limited, the mahoganoid gene product.

Once isolated, such a protein can be identified and can be used in conjunction with standard techniques, to identify proteins it interacts with. For example, at least a portion of the amino acid sequence of a protein that interacts with the mahogany gene product can be ascertained using techniques well known to those of skill in the art, such as via the Edman degradation technique (see, *e.g.*, Creighton, 1983, "Proteins: Structures and Molecular Principles," W.H. Freeman & Co., N.Y., pp.34-49). The amino acid sequence obtained may be used as a guide for the generation of oligonucleotide mixtures that can be used to screen for gene sequences encoding such proteins. Screening may be accomplished, for example, by standard hybridization or PCR techniques. Techniques for the generation of oligonucleotide mixtures and the screening are well-known. (See, *e.g.*, Ausubel, *supra*, and 1990, "PCR Protocols: A Guide to Methods and Applications," Innis, *et al.*, eds. Academic Press, Inc., New York).

Additionally, methods may be employed that result in the simultaneous identification of genes that encode a protein which interacts with a mahogany gene product. These methods include, for example, probing expression libraries with

labeled mahogany gene product, using mahogany gene product in a manner similar to the well known technique of antibody probing of λgt11 libraries.

One method that detects protein interactions *in vivo*, the two-hybrid system, is described in detail for illustration only and not by way of limitation. One version of this system has been described (Chien, *et al.*, 1991, Proc. Natl. Acad. Sci. USA, 88, 9578-9582) and is commercially available from Clontech (Palo Alto, CA).

Briefly, utilizing such a system, plasmids are constructed that encode two hybrid proteins: one consists of the DNA-binding domain of a transcription activator protein fused to the mahogany gene product and the other consists of the transcription activator protein's activation domain fused to an unknown protein that is encoded by a cDNA that has been recombined into this plasmid as part of a cDNA library. The DNA-binding domain fusion plasmid and the cDNA library are transformed into a strain of the yeast *Saccharomyces cerevisiae* that contains a reporter gene (*e.g.*, HBS or *lacZ*) whose regulatory region contains the transcription activator's binding site. Either hybrid protein alone cannot activate transcription of the reporter gene: the DNA-binding domain hybrid cannot because it does not provide activation function and the activation domain hybrid cannot because it cannot localize to the activator's binding sites. Interaction of the two hybrid proteins reconstitutes the functional activator protein and results in expression of the reporter gene, which is detected by an assay for the reporter gene product.

The two-hybrid system or related methodologies may be used to screen activation domain libraries for proteins that interact with the "bait" gene product. By way of example, and not by way of limitation, mahogany gene products may be used as the bait gene product. Total genomic or cDNA

sequences are fused to the DNA encoding an activation domain. This library and a plasmid encoding a hybrid of a bait mahogany gene product fused to the DNA-binding domain are co-transformed into a yeast reporter strain, and the resulting transformants are screened for those that express the reporter gene. For example, a bait mahogany gene sequence, such as the open reading frame of the mahogany gene, can be cloned into a vector such that it is translationally fused to the DNA encoding the DNA-binding domain of the GAL4 protein. These colonies are purified and the library plasmids responsible for reporter gene expression are isolated. DNA sequencing is then used to identify the proteins encoded by the library plasmids.

A cDNA library of the cell line from which proteins that interact with bait mahogany gene product are to be detected can be made using methods routinely practiced in the art. According to the particular system described herein, for example, the cDNA fragments can be inserted into a vector such that they are translationally fused to the transcriptional activation domain of GAL4. Such a library can be co-transformed along with the bait mahogany gene-GAL4 fusion plasmid into a yeast strain that contains a lacZ gene driven by a promoter that contains GAL4 activation sequence. A cDNA encoded protein, fused to a GAL4 transcriptional activation domain that interacts with bait mahogany gene product will reconstitute an active GAL4 protein and thereby drive expression of the HIS3 gene. Colonies that express HIS3 can be detected by their growth on petri dishes containing semi-solid agar based media lacking histidine. The cDNA can then be purified from these strains, and used to produce and isolate the bait mahogany gene product-interacting protein using techniques routinely practiced in the art.

5.4.2.3. ASSAYS FOR COMPOUNDS THAT INTERFERE WITH MAHOGANY GENE PRODUCT MACROMOLECULE INTERACTION

The mahogany gene products may, *in vivo*, interact with one or more macromolecules, such as proteins. For example, the mahogany gene products may, *in vivo*, interact with the mahoganoid gene products. Other macromolecules which interact with the mahogany gene products may include, but are not limited to, nucleic acid molecules and those proteins identified via methods such as those described, above, in Sections 5.4.2.1 - 5.4.2.2. For purposes of this discussion, the macromolecules are referred to herein as "binding partners". Compounds that disrupt mahogany gene product binding to a binding partner may be useful in regulating the activity of the mahogany gene product, especially mutant mahogany gene products. Such compounds may include, but are not limited to molecules such as peptides, and the like, as described, for example, in Section 5.4.2.1 above.

The basic principle of an assay system used to identify compounds that interfere with the interaction between the mahogany gene product and a binding partner or partners involves preparing a reaction mixture containing the mahogany gene product and the binding partner under conditions and for a time sufficient to allow the two to interact and bind, thus forming a complex. In order to test a compound for inhibitory activity, the reaction mixture is prepared in the presence and absence of the test compound. The test compound may be initially included in the reaction mixture, or may be added at a time subsequent to the addition of mahogany gene product and its binding partner. Control reaction mixtures are incubated without the test compound or with a compound which is known not to block complex formation. The formation of any complexes between the mahogany gene product and the binding partner is then detected. The formation of a complex

in the control reaction, but not in the reaction mixture containing the test compound, indicates that the compound interferes with the interaction of the mahogany gene product and the binding partner. Additionally, complex formation within reaction mixtures containing the test compound and normal mahogany gene product may also be compared to complex formation within reaction mixtures containing the test compound and a mutant mahogany gene product. This comparison may be important in those cases wherein it is desirable to identify compounds that disrupt interactions of mutant but not normal mahogany gene product.

The assay for compounds that interfere with the interaction of the mahogany gene products and binding partners can be conducted in a heterogeneous or homogeneous format. Heterogeneous assays involve anchoring either the mahogany gene product or the binding partner onto a solid support and detecting complexes formed on the solid support at the end of the reaction. In homogeneous assays, the entire reaction is carried out in a liquid phase. In either approach, the order of addition of reactants can be varied to obtain different information about the compounds being tested. For example, test compounds that interfere with the interaction between the mahogany gene products and the binding partners, *e.g.*, by competition, can be identified by conducting the reaction in the presence of the test substance; *i.e.*, by adding the test substance to the reaction mixture prior to or simultaneously with the mahogany gene product and interactive intracellular binding partner. Alternatively, test compounds that disrupt preformed complexes, *e.g.*, compounds with higher binding constants that displace one of the components from the complex, can be tested by adding the test compound to the reaction mixture after complexes have been formed. The various formats are described briefly below.

In a heterogeneous assay system, either the mahogany gene product or the interactive binding partner, is anchored onto a solid surface, while the non-anchored species is labeled, either directly or indirectly. In practice, microtiter plates are conveniently utilized. The anchored species may be immobilized by non-covalent or covalent attachments. Non-covalent attachment may be accomplished simply by coating the solid surface with a solution of the mahogany gene product or binding partner and drying. Alternatively, an immobilized antibody specific for the species to be anchored may be used to anchor the species to the solid surface. The surfaces may be prepared in advance and stored.

In order to conduct the assay, the partner of the immobilized species is exposed to the coated surface with or without the test compound. After the reaction is complete, unreacted components are removed (*e.g.*, by washing) and any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of ways. Where the non-immobilized species is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the non-immobilized species is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface; *e.g.*, using a labeled antibody specific for the initially non-immobilized species (the antibody, in turn, may be directly labeled or indirectly labeled with a labeled anti-Ig antibody). Depending upon the order of addition of reaction components, test compounds that inhibit complex formation or that disrupt preformed complexes can be detected.

Alternatively, the reaction can be conducted in a liquid phase in the presence or absence of the test compound, the reaction products separated from unreacted components, and

complexes detected; *e.g.*, using an immobilized antibody specific for one of the binding components to anchor any complexes formed in solution, and a labeled antibody specific for the other partner to detect anchored complexes. Again, depending upon the order of addition of reactants to the liquid phase, test compounds that inhibit complex formation or that disrupt preformed complexes can be identified.

In an alternate embodiment of the invention, a homogeneous assay can be used. In this approach, a preformed complex of the mahogany gene product and the interactive binding partner is prepared in which either the mahogany gene product or its binding partners is labeled, but the signal generated by the label is quenched due to complex formation (see, *e.g.*, U.S. Patent No. 4,109,496 by Rubenstein which utilizes this approach for immunoassays). The addition of a test substance that competes with and displaces one of the species from the preformed complex will result in the generation of a signal above background. In this way, test substances that disrupt mahogany gene product/binding partner interaction can be identified.

In another embodiment of the invention, these same techniques can be employed using peptide fragments that correspond to the binding domains of the mahogany gene product and/or the binding partner (in cases where the binding partner is a protein), in place of one or both of the full length proteins. Any number of methods routinely practiced in the art can be used to identify and isolate the binding sites. These methods include, but are not limited to, mutagenesis of the gene encoding one of the proteins and screening for disruption of binding in a co-immunoprecipitation assay. Compensating mutations in the gene encoding the second species in the complex can then be selected. Sequence analysis of the genes encoding the respective proteins will reveal the mutations that correspond

to the region of the protein involved in interactive binding. Alternatively, one protein can be anchored to a solid surface using methods described in this Section above, and allowed to interact with and bind to its labeled binding partner, which has been treated with a proteolytic enzyme, such as trypsin. After washing, a short, labeled peptide comprising the binding domain may remain associated with the solid material, which can be isolated and identified by amino acid sequencing. Also, once the gene coding for the segments is engineered to express peptide fragments of the protein, it can then be tested for binding activity and purified or synthesized.

For example, and not by way of limitation, a mahogany gene product can be anchored to a solid material as described, above, in this Section by making a GST-1 fusion protein and allowing it to bind to glutathione agarose beads. The binding partner can be labeled with a radioactive isotope, such as ^{35}S, and cleaved with a proteolytic enzyme such as trypsin. Cleavage products can then be added to the anchored GST-1 fusion protein and allowed to bind. After washing away unbound peptides, labeled bound material, representing the binding partner binding domain, can be eluted, purified, and analyzed for amino acid sequence by well-known methods. Peptides so identified can be produced synthetically or produced using recombinant DNA technology.

5.4.2.4. ASSAYS FOR THE IDENTIFICATION OF COMPOUNDS THAT AMELIORATE BODY WEIGHT DISORDERS

Compounds, including but not limited to binding compounds identified via assay techniques such as those described, above, in Sections 5.4.2.1 - 5.4.2.3, can be tested for the ability to ameliorate body weight disorder symptoms, including obesity, cachexia, and anorexia. It

should be noted that the assays described herein can identify compounds that affect mahogany activity by either affecting mahogany gene expression or by affecting the level of mahogany gene product activity. For example, compounds may be identified that are involved in another step in the pathway in which the mahogany gene and/or mahogany gene product is involved, such as, for example, a step which is either "upfield" or "downfield" of the step in the pathway mediated by the mahogany gene. Such compounds may, by affecting this same pathway, modulate the effect of mahogany on the development of body weight disorders. Such compounds can be used as part of a therapeutic method for the treatment of the disorder.

Described below are cell-based and animal model-based assays for the identification of compounds exhibiting such an ability to ameliorate body weight disorder symptoms.

First, cell-based systems can be used to identify compounds that may act to ameliorate body weight disorder symptoms. Such cell systems can include, for example, recombinant or non-recombinant cell, such as cell lines, that express the mahogany gene.

In utilizing such cell systems, cells that express mahogany may be exposed to a compound suspected of exhibiting an ability to ameliorate body weight disorder symptoms, at a sufficient concentration and for a sufficient time to elicit such an amelioration of such symptoms in the exposed cells. After exposure, the cells can be assayed to measure alterations in the expression of the mahogany gene, *e.g.*, by assaying cell lysates for mahogany mRNA transcripts (*e.g.*, by Northern analysis) or for mahogany gene products expressed by the cell; compounds that modulate expression of the mahogany gene are good candidates as therapeutics.

In addition, animal-based systems or models for a mammalian body weight disorder, for example, transgenic mice

containing a human or altered form of mahogany gene, may be used to identify compounds capable of ameliorating symptoms of the disorder. Such animal models may be used as test substrates for the identification of drugs, pharmaceuticals, therapies and interventions. For example, animal models may be exposed to a compound suspected of exhibiting an ability to ameliorate symptoms, at a sufficient concentration and for a sufficient time to elicit such an amelioration of body weight disorder symptoms. The response of the animals to the exposure may be monitored by assessing the reversal of the symptoms of the disorder.

With regard to intervention, any treatments that reverse any aspect of body weight disorder-like symptoms should be considered as candidates for human therapeutic intervention in such a disorder. Dosages of test agents may be determined by deriving dose-response curves, as discussed in Section 5.5.1, below.

5.4.3. COMPOUNDS AND METHODS FOR THE TREATMENT OF BODY WEIGHT DISORDERS

Described below are methods and compositions whereby body weight disorders, including obesity, cachexia, and anorexia, may be treated. Such methods can comprise, for example administering compounds which modulate the expression of a mammalian mahogany gene and/or the synthesis or activity of a mammalian mahogany gene product, so that symptoms of the body weight disorder are ameliorated. Alternatively, in those instances whereby the mammalian body weight disorder results from mahogany gene mutations, such methods can comprise supplying the mammal with a nucleic acid molecule encoding an unimpaired mahogany gene product such that an unimpaired mahogany gene product is expressed and symptoms of the disorder are ameliorated.

In another embodiment of methods for the treatment of mammalian body weight disorders resulting from mahogany gene mutations, such methods can comprise supplying the mammal with a cell comprising a nucleic acid molecule that encodes an unimpaired mahogany gene product such that the cell expresses the unimpaired mahogany gene product, and symptoms of the disorder are ameliorated.

Because a loss of normal mahogany gene function results in the restoration of a non-obese phenotype in individuals exhibiting an agouti mutation (*e.g.* individuals that ectopically express the agouti gene in all tissues) a decrease or elimination of normal mahogany gene product would facilitate progress towards a normal body weight state in such individuals. Methods for inhibiting or reducing the level of mahogany gene product synthesis or expression can include, for example, methods such as those described in Section 5.4.3.1.

Alternatively, symptoms of certain body weight disorders such as, for example, cachexia and anorexia, which involve a lower than normal body weight phenotype, may be ameliorated by increasing the level of mahogany gene expression and/or mahogany gene product activity. Methods for enhancing the expression or synthesis of mahogany can include, for example, methods such as those described below, in Section 5.4.3.2

5.4.3.1. INHIBITORY ANTISENSE, RIBOZYME AND TRIPLE HELIX APPROACHES

In another embodiment, symptoms of body weight disorders may be ameliorated by decreasing the level of mahogany gene expression and/or mahogany gene product activity by using mahogany gene sequences in conjunction with well-known antisense, gene "knock-out," ribozyme and/or triple helix methods to decrease the level of mahogany gene expression. Among the compounds that may exhibit the ability to modulate

the activity, expression or synthesis of the mahogany gene, including the ability to ameliorate the symptoms of a mammalian body weight disorder, are antisense, ribozyme, and triple helix molecules. Such molecules may be designed to reduce or inhibit either unimpaired, or if appropriate, mutant target gene activity. Techniques for the production and use of such molecules are well known to those of skill in the art.

Antisense RNA and DNA molecules act to directly block the translation of mRNA by hybridizing to targeted mRNA and preventing protein translation. Antisense approaches involve the design of oligonucleotides that are complementary to a target gene mRNA. The antisense oligonucleotides will bind to the complementary target gene mRNA transcripts and prevent translation. Absolute complementarily, although preferred, is not required.

A sequence "complementary" to a portion of an RNA, as referred to herein, means a sequence having sufficient complementarily to be able to hybridize with the RNA, forming a stable duplex; in the case of double-stranded antisense nucleic acids, a single strand of the duplex DNA may thus be tested, or triplex formation may be assayed. The ability to hybridize will depend on both the degree of complementarily and the length of the antisense nucleic acid. Generally, the longer the hybridizing nucleic acid, the more base mismatches with an RNA it may contain and still form a stable duplex (or triplex, as the case may be). One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

In one embodiment, oligonucleotides complementary to non-coding regions of the mahogany gene could be used in an antisense approach to inhibit translation of endogenous mahogany mRNA. Antisense nucleic acids should be at least

six nucleotides in length, and are preferably oligonucleotides ranging from 6 to about 50 nucleotides in length. In specific aspects the oligonucleotide is at least 10 nucleotides, at least 17 nucleotides, at least 25 nucleotides or at least 50 nucleotides.

Regardless of the choice of target sequence, it is preferred that *in vitro* studies are first performed to quantitate the ability of the antisense oligonucleotide to inhibit gene expression. It is preferred that these studies utilize controls that distinguish between antisense gene inhibition and nonspecific biological effects of oligonucleotides. It is also preferred that these studies compare levels of the target RNA or protein with that of an internal control RNA or protein. Additionally, it is envisioned that results obtained using the antisense oligonucleotide are compared with those obtained using a control oligonucleotide. It is preferred that the control oligonucleotide is of approximately the same length as the test oligonucleotide and that the nucleotide sequence of the oligonucleotide differs from the antisense sequence no more than is necessary to prevent specific hybridization to the target sequence.

The oligonucleotides can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridization, etc. The oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger, *et al.*, 1989, Proc. Natl. Acad. Sci. U.S.A. 86, 6553-6556; Lemaitre, *et al.*, 1987, Proc. Natl. Acad. Sci. U.S.A. 84, 648-652; PCT Publication No. WO88/09810, published December 15, 1988) or

the blood-brain barrier (see, *e.g.*, PCT Publication No. WO89/10134, published April 25, 1988), hybridization-triggered cleavage agents (see, *e.g.*, Krol *et al.*, 1988, BioTechniques 6, 958-976) or intercalating agents (see, *e.g.*, Zon, 1988, Pharm. Res. 5, 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

The antisense oligonucleotide may comprise at least one modified base moiety which is selected from the group including but not limited to 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

The antisense oligonucleotide may also comprise at least one modified sugar moiety selected from the group including but not limited to arabinose, 2-fluoroarabinose, xylulose, and hexose.

In yet another embodiment, the antisense oligonucleotide comprises at least one modified phosphate backbone selected

from the group consisting of a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

In yet another embodiment, the antisense oligonucleotide is an α-anomeric oligonucleotide. An α-anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gautier, *et al.*, 1987, Nucl. Acids Res. 15, 6625-6641). The oligonucleotide is a 2'-0-methylribonucleotide (Inoue, *et al.*, 1987, Nucl. Acids Res. 15, 6131-6148), or a chimeric RNA-DNA analogue (Inoue, *et al.*, 1987, FEBS Lett. 215, 327-330).

Oligonucleotides of the invention may be synthesized by standard methods known in the art, *e.g.*, by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides may be synthesized by the method of Stein, *et al.* (1988, Nucl. Acids Res. 16, 3209), methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin, *et al.*, 1988, Proc. Natl. Acad. Sci. U.S.A. 85, 7448-7451), etc.

While antisense nucleotides complementary to the target gene coding region sequence could be used, those complementary to the transcribed, untranslated region are most preferred.

Antisense molecules should be delivered to cells that express the target gene *in vivo*. A number of methods have been developed for delivering antisense DNA or RNA to cells; *e.g.*, antisense molecules can be injected directly into the tissue site, or modified antisense molecules, designed to target the desired cells (*e.g.*, antisense linked to peptides or antibodies that specifically bind receptors or antigens

expressed on the target cell surface) can be administered systemically.

However, it is often difficult to achieve intracellular concentrations of the antisense sufficient to suppress translation of endogenous mRNAs. Therefore a preferred approach utilizes a recombinant DNA construct in which the antisense oligonucleotide is placed under the control of a strong pol III or pol II promoter. The use of such a construct to transfect target cells in the patient will result in the transcription of sufficient amounts of single stranded RNAs that will form complementary base pairs with the endogenous target gene transcripts and thereby prevent translation of the target gene mRNA. For example, a vector can be introduced *e.g.*, such that it is taken up by a cell and directs the transcription of an antisense RNA. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA. Such vectors can be constructed by recombinant DNA technology methods standard in the art. Vectors can be plasmid, viral, or others known in the art, used for replication and expression in mammalian cells. Expression of the sequence encoding the antisense RNA can be by any promoter known in the art to act in mammalian, preferably human cells. Such promoters can be inducible or constitutive. Such promoters include but are not limited to: the SV40 early promoter region (Bernoist and Chambon, 1981, Nature 290, 304-310), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto, *et al.*, 1980, Cell 22, 787-797), the herpes thymidine kinase promoter (Wagner, *et al.*, 1981, Proc. Natl. Acad. Sci. U.S.A. 78, 1441-1445), the regulatory sequences of the metallothionein gene (Brinster, *et al.*, 1982, Nature 296, 39-42), etc. Any type of plasmid, cosmid, YAC or viral vector can be used to prepare the recombinant DNA construct which can be introduced

directly into the tissue site. Alternatively, viral vectors can be used that selectively infect the desired tissue, in which case administration may be accomplished by another route (*e.g.*, systemically).

Ribozyme molecules designed to catalytically cleave target gene mRNA transcripts can also be used to prevent translation of target gene mRNA and, therefore, expression of target gene product. (See, *e.g.*, PCT International Publication WO90/11364, published October 4, 1990; Sarver, *et al.*, 1990, Science 247, 1222-1225).

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. (For a review, see Rossi, 1994, Current Biology 4, 469-471). The mechanism of ribozyme action involves sequence specific hybridization of the ribozyme molecule to complementary target RNA, followed by an endonucleolytic cleavage event. The composition of ribozyme molecules must include one or more sequences complementary to the target gene mRNA, and must include the well known catalytic sequence responsible for mRNA cleavage. For this sequence, see, *e.g.*, U.S. Patent No. 5,093,246, which is incorporated herein by reference in its entirety.

While ribozymes that cleave mRNA at site specific recognition sequences can be used to destroy target gene mRNAs, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA have the following sequence of two bases: 5'-UG-3'. The construction and production of hammerhead ribozymes is well known in the art and is described more fully in Myers, 1995, *Molecular Biology and Biotechnology: A Comprehensive Desk Reference*, VCH Publishers, New York, (see especially Figure 4, page 833) and in Haseloff and Gerlach, 1988, Nature, 334,

585-591, which is incorporated herein by reference in its entirety.

Preferably the ribozyme is engineered so that the cleavage recognition site is located near the 5' end of the target gene mRNA, *i.e.*, to increase efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts.

The ribozymes of the present invention also include RNA endoribonucleases (hereinafter "Cech-type ribozymes") such as the one that occurs naturally in *Tetrahymena thermophila* (known as the IVS, or L-19 IVS RNA) and that has been extensively described by Thomas Cech and collaborators (Zaug, *et al.*, 1984, Science, 224, 574-578; Zaug and Cech, 1986, Science, 231, 470-475; Zaug, *et al.*, 1986, Nature, 324, 429-433; published International patent application No. WO 88/04300 by University Patents Inc.; Been and Cech, 1986, Cell, 47, 207-216). The Cech-type ribozymes have an eight base pair active site which hybridizes to a target RNA sequence whereafter cleavage of the target RNA takes place. The invention encompasses those Cech-type ribozymes which target eight base-pair active site sequences that are present in the target gene.

As in the antisense approach, the ribozymes can be composed of modified oligonucleotides (*e.g.*, for improved stability, targeting, etc.) and should be delivered to cells that express the target gene *in vivo*. A preferred method of delivery involves using a DNA construct "encoding" the ribozyme under the control of a strong constitutive pol III or pol II promoter, so that transfected cells will produce sufficient quantities of the ribozyme to destroy endogenous target gene messages and inhibit translation. Because ribozymes unlike antisense molecules, are catalytic, a lower intracellular concentration is required for efficiency.

Endogenous target gene expression can also be reduced by inactivating or "knocking out" the target gene or its promoter using targeted homologous recombination (*e.g.*, see Smithies, *et al.*, 1985, Nature 317, 230-234; Thomas and Capecchi, 1987, Cell 51, 503-512; Thompson, *et al.*, 1989, Cell 5, 313-321; each of which is incorporated by reference herein in its entirety). For example, a mutant, non-functional target gene (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous target gene (either the coding regions or regulatory regions of the target gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express the target gene *in vivo*. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the target gene. Such approaches are particularly suited in the agricultural field where modifications to ES (embryonic stem) cells can be used to generate animal offspring with an inactive target gene (*e.g.*, see Thomas and Capecchi, 1987 and Thompson, 1989, *supra*). However this approach can be adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site *in vivo* using appropriate viral vectors.

Alternatively, endogenous target gene expression can be reduced by targeting deoxyribonucleotide sequences complementary to the regulatory region of the target gene (*i.e.*, the target gene promoter and/or enhancers) to form triple helical structures that prevent transcription of the target gene in target cells in the body. (See generally, Helene, 1991, Anticancer Drug Des., 6(6), 569-584; Helene, *et al.*, 1992, Ann. N.Y. Acad. Sci., 660, 27-36; and Maher, 1992, Bioassays 14(12), 807-815).

Nucleic acid molecules to be used in triplex helix formation for the inhibition of transcription should be

single stranded and composed of deoxynucleotides. The base composition of these oligonucleotides must be designed to promote triple helix formation via Hoogsteen base pairing rules, which generally require sizeable stretches of either purines or pyrimidines to be present on one strand of a duplex. Nucleotide sequences may be pyrimidine-based, which will result in TAT and CGC+ triplets across the three associated strands of the resulting triple helix. The pyrimidine-rich molecules provide base complementarily to a purine-rich region of a single strand of the duplex in a parallel orientation to that strand. In addition, nucleic acid molecules may be chosen that are purine-rich, for example, contain a stretch of G residues. These molecules will form a triple helix with a DNA duplex that is rich in GC pairs, in which the majority of the purine residues are located on a single strand of the targeted duplex, resulting in GGC triplets across the three strands in the triplex.

Alternatively, the potential sequences that can be targeted for triple helix formation may be increased by creating a so called "switchback" nucleic acid molecule. Switchback molecules are synthesized in an alternating 5'-3', 3'-5' manner, such that they base pair with first one strand of a duplex and then the other, eliminating the necessity for a sizeable stretch of either purines or pyrimidines to be present on one strand of a duplex.

In instances wherein the antisense, ribozyme, and/or triple helix molecules described herein are utilized to inhibit mutant gene expression, it is possible that the technique may so efficiently reduce or inhibit the transcription (triple helix) and/or translation (antisense, ribozyme) of mRNA produced by normal target gene alleles that the possibility may arise wherein the concentration of normal target gene product present may be lower than is necessary for a normal phenotype. In such cases, to ensure that

substantially normal levels of target gene activity are maintained, therefore, nucleic acid molecules that encode and express target gene polypeptides exhibiting normal target gene activity may, be introduced into cells via gene therapy methods such as those described, below, in Section 5.9.2 that do not contain sequences susceptible to whatever antisense, ribozyme, or triple helix treatments are being utilized. Alternatively, in instances whereby the target gene encodes an extracellular protein, it may be preferable to co-administer normal target gene protein in order to maintain the requisite level of target gene activity.

Anti-sense RNA and DNA, ribozyme, and triple helix molecules of the invention may be prepared by any method known in the art for the synthesis of DNA and RNA molecules, as discussed above. These include techniques for chemically synthesizing oligodeoxyribonucleotides and oligoribonucleotides well known in the art such as for example solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be incorporated into a wide variety of vectors that incorporate suitable RNA polymerase promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

5.4.3.2. GENE REPLACEMENT THERAPY

Mahogany gene nucleic acid sequences, described above in Section 5.1, can be utilized for the treatment of a mammalian body weight disorders, including obesity, cachexia, and anorexia. Such treatment can be in the form of gene replacement therapy. Specifically, one or more copies of a

normal mahogany gene or a portion of the mahogany gene that directs the production of a mahogany gene product exhibiting normal mahogany gene function, may be inserted into the appropriate cells within a patient, using vectors that include, but are not limited to adenovirus, adeno-associated virus, and retrovirus vectors, in addition to other particles that introduce DNA into cells, such as liposomes.

Because the mahogany gene is expressed in the brain, such gene replacement therapy techniques should be capable delivering mahogany gene sequences to these cell types within patients. Thus, in one embodiment, techniques that are well known to those of skill in the art (see, *e.g.*, PCT Publication No. WO89/10134, published April 25, 1988) can be used to enable mahogany gene sequences to cross the blood-brain barrier readily and to deliver the sequences to cells in the brain. With respect to delivery that is capable of crossing the blood-brain barrier, viral vectors such as, for example, those described above, are preferable.

In another embodiment, techniques for delivery involve direct administration of such mahogany gene sequences to the site of the cells in which the mahogany gene sequences are to be expressed.

Additional methods that may be utilized to increase the overall level of mahogany gene expression and/or mahogany gene product activity include using target homologous recombination methods, discussed in Section 5.2, above, to modify the expression characteristic of an endogenous mahogany gene in a cell or microorganism by inserting a heterologous DNA regulatory element such that the inserted regulatory element is operatively linked with the endogenous mahogany gene in question. Targeted homologous recombination can be thus used to activated transcription of an endogenous mahogany gene that is "transcriptionally silent", *i.e.*, is

not normally expressed, or to enhance the expression of an endogenous mahogany gene that is normally expressed.

Further, the overall level of mahogany gene expression and/or mahogany gene product activity may be increased by the introduction of appropriate mahogany-expressing cells, preferably autologous cells, into a patient at positions and in numbers that are sufficient to ameliorate body weight disorder symptoms. Such cells may be either recombinant or non-recombinant.

Among the cells that can be administered to increase the overall level of mahogany gene expression in a patient are normal cells, preferably brain cells, that express the mahogany gene. Alternatively, cells, preferably autologous cells, can be engineered to express mahogany gene sequences, and may then be introduced into a patient in positions appropriate for the amelioration of the body weight disorder symptoms. Alternately, cells that express an unimpaired mahogany gene and that are from a MHC matched individual can be utilized, and may include, for example, brain cells. The expression of the mahogany gene sequences is controlled by the appropriate gene regulatory sequences to allow such expression in the necessary cell types. Such gene regulatory sequences are well known to the skilled artisan. Such cell-based gene therapy techniques are well known to those skilled in the art, see, *e.g.*, Anderson, U.S. Patent No. 5,399,349.

When the cells to be administered are non-autologous cells, they can be administered using well known techniques that prevent a host immune response against the introduced cells from developing. For example, the cells may be introduced in an encapsulated form which, while allowing for an exchange of components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

Additionally, compounds, such as those identified via techniques such as those described, above, in Section 5.4.2, that are capable of modulating mahogany gene product activity can be administered using standard techniques that are well known to those of skill in the art. In instances in which the compounds to be administered are to involve an interaction with brain cells, the administration techniques should include well known ones that allow for a crossing of the blood-brain barrier.

5.5. PHARMACEUTICAL PREPARATIONS AND METHODS OF ADMINISTRATION

The compounds that are determined to affect mahogany gene expression or gene product activity can be administered to a patient at therapeutically effective doses to treat or ameliorate body weight disorders, such as obesity, anorexia, or cachexia. A therapeutically effective dose refers to that amount of the compound sufficient to result in amelioration of symptoms of such a disorder.

5.5.1. EFFECTIVE DOSE

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD_{50}/ED_{50}. Compounds that exhibit large therapeutic indices are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC_{50} (*i.e.*, the concentration of the test compound that achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

5.5.2. FORMULATIONS AND USE

Pharmaceutical compositions for use in accordance with the present invention may be formulated in conventional manner using one or more physiologically acceptable carriers or excipients.

Thus, the compounds and their physiologically acceptable salts and solvates may be formulated for administration by inhalation or insufflation (either through the mouth or the nose) or oral, buccal, parenteral or rectal administration.

For oral administration, the pharmaceutical compositions may take the form of, for example, tablets or capsules prepared by conventional means with pharmaceutically acceptable excipients such as binding agents (*e.g.*, pregelatinised maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose); fillers (*e.g.*, lactose, microcrystalline cellulose or calcium hydrogen phosphate);

lubricants (*e.g.*, magnesium stearate, talc or silica); disintegrants (*e.g.*, potato starch or sodium starch glycolate); or wetting agents (*e.g.*, sodium lauryl sulphate). The tablets may be coated by methods well known in the art. Liquid preparations for oral administration may take the form of, for example, solutions, syrups or suspensions, or they may be presented as a dry product for constitution with water or other suitable vehicle before use. Such liquid preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (*e.g.*, sorbitol syrup, cellulose derivatives or hydrogenated edible fats); emulsifying agents (*e.g.*, lecithin or acacia); non-aqueous vehicles (*e.g.*, almond oil, oily esters, ethyl alcohol or fractionated vegetable oils); and preservatives (*e.g.*, methyl or propyl-p-hydroxybenzoates or sorbic acid). The preparations may also contain buffer salts, flavoring, coloring and sweetening agents as appropriate.

Preparations for oral administration may be suitably formulated to give controlled release of the active compound.

For buccal administration the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebulizer, with the use of a suitable propellant, *e.g.*, dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of *e.g.*, gelatin for use in an inhaler or insufflator may be formulated containing a

powder mix of the compound and a suitable powder base such as lactose or starch.

The compounds may be formulated for parenteral administration by injection, *e.g.*, by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, *e.g.*, in ampoules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, *e.g.*, sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, *e.g.*, containing conventional suppository bases such as cocoa butter or other glycerides.

In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

The compositions may, if desired, be presented in a pack or dispenser device that may contain one or more unit dosage forms containing the active ingredient. The pack may for example comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration.

6. EXAMPLE: GENETIC AND PHYSICAL MAPPING OF THE MAHOGANY LOCUS

In the Example presented herein, studies are described which, first, define the genetic interval on mouse chromosome 2 within which the mahogany gene lies, and second, successfully narrow the interval to approximately 0.29 cM. Further, the physical mapping of this interval is described.

Mouse crosses were performed to obtain homozygous mg/mg mice. First, LDJ-Le-mg mice were crossed with CAST/Ei mice. The F1s were back-crossed with LDJ-Le-mg mice and the resulting litters scored for coat color. Mice showing coat color of mg/mg homozygotes were genotyped to using D2/NDS3 and D2/MIT19 markers to identify meiotic events. Mice showing recombinant events were fine structure mapped using various markers shown in FIG. 1. All genotyping was performed using PCR-SSLP and then analyzed using PAGE.

After 2300 meiosesis, the mahogany gene was mapped to a 0.99 cM interval FIG. 1. This corresponded to an interval width of 700 kb.

Physical Mapping of the Genetic Interval: The 700 kb mahogany region on mouse chromosome 2 is shown in FIG. 1. Genetic markers, clones spanning the region and open reading frames in the interval are shown in the figure.

7. EXAMPLE: IDENTIFICATION OF A CANDIDATE MAHOGANY GENE

In the Example presented herein, a gene is identified within the cloned DNA described in the Example in Section 6, above, which corresponds to a candidate mahogany gene.

Clones spanning the 700kb region were sequenced and open reading frames were identified and analyzed through this interval. Nucleic acid sequencing was performed using ABI sequencers and the manufactures recommended procedures. Many

novel sequences encoding proteins are located in this integral, see the bottom of FIG. 1. With each open reading frame identified, mutational analysis, primarily via SSCP analysis, was used with the three alleles of the mahogany phenotype mice to identify which of the open reading frames within this interval contain a mutation in an mg mouse.

A mutation was found in one of the genomic/cDNA sequences found in the integral in mg3J mice. Figures 3 and 2 provide the genomic and cDNA sequences surrounding the mutation, FIG. 6 shows the mutation in mg3J, and FIGS. 8 and 9 show splice variants in the 5' end of the murine mg gene. The mutation in mg3J mice is a deletion of a GCTGC sequence which results in the creation of a frameshift. Based on the chromosomal location and mutation identification, the cDNA provided in Figure 2 and the corresponding genomic DNA which contains the contigs provided in Figure 3 represent the mg gene/locus.

Further analysis of cDNA clones identified two distinct splice variants in the 5' end of the mg gene. Figure 7 provides an analysis of the structure of the two splice variants, denoted akm1003 and akm1004. Figures 8 and 9 provide the nucleic acid and amino acid sequence of the 5' ends of these splice variants and structural analysis of the protein encoded by the 5' regions.

Analysis of libraries of human cDNA sequences led to the identification of three forms of the human ortholog of the mg gene: a long form (FIG. 18) and two shorter splice forms, each of which is shown in FIGS. 19 and 20.

8. EXAMPLE: CHARACTERIZATION OF THE MAHOGANY GENE

In the example presented herein, the nucleic acid sequence of the mahogany gene transcript identified in the example presented in Section 7, above, is used to generate

Northern analysis data which characterize the expression of the mahogany transcript in a number of tissues both of wild type mice, and of mice exhibiting the mahogany phenotype. The results presented in this example are consistent with the mg gene being the mahogany gene.

For Northern analysis, polyA RNA was isolated from wild-type and the original mg mutant, mg3J and mg-Lester mice and utilized from the Northern analysis following standard protocols. Northern blots prepared from this mRNA was hybridized with a probe obtained from sequences common to the akm1003 and akm1004 sequences. Specifically PCR primers TTCCTCACTGG and GGACACACAG were used to amplify cDNA from the akm1003 sequence which had been radiolabelled by random priming using a Gibco-BRL kit according to the manufacturer's recommended protocol.

An mg transcript was found in all mice examined in mRNA isolated from brain (minus the hypothalamus), kidney, heart, testes, liver, skin, and hypothalamus. No expression was seen in muscle.

In a Northern blot run on RNA samples from mahogany mice, the mg transcript was found to be expressed at a reduced level in all tissues in mRNA isolated from mg3J mice, as a varied size fragment in mg-Lester derived mRNA, and at different levels and sizes in original mg mutant mice derived mRNA.

These results are consistent with the mg gene disclosed herein as being the mahogany gene.

9. EXAMPLE: EFFECTS OF THE MAHOGANY GENE ON GENETIC AND DIETARY OBESITY

This section describes experiments which examine whether the *mg* gene acts specifically within the agouti pathway. Specifically, these experiments test whether *mg* can suppress the obesity of other monogenic obese mutants as well as

whether it can suppress diet-induced obesity. The results show that *mg* does not suppress obesity in any of the monogenic obese mutants. However, *mg* can suppress diet-induced obesity. Thus, the *mg* gene and its corresponding gene product and compounds that modulate *mg* expression and/or activity have implications in the treatment of diet- induced obesity disorders, as well as in the treatment of disorders related directly to the *mg* or agouti gene.

9.1. MATERIALS AND METHODS

Genetic crosses: The crosses, and the number of animals for each (n) were (LDJ/Le-*mg/mg* X CAST/Ei) X LDJ/Le-*mg/mg* (n=1588), (C3HeB/FeJ-mg^{3J}/ mg^{3J} X CAST/Ei) X C3HeB/FeJ-mg^{3J}/ mg^{3J} (n=324), (C3HeB/FeJ-mg^{3J}/ mg^{3J} X MOLF/Ei) X C3HeB/FeJ-mg^{3J}/ mg^{3J} (n=216) and (C3HeB/FeJ-mg^{3J}/ mg^{3J} X C57BL6/J) X C3HeB/FeJ-mg^{3J}/ mg^{3J} (n=309). The 2437 N_2 mice were analysed by coat colour to determine their genotype at the *mg* locus. As mice change color slightly at each hair molt and because the phenotype of *mg/mg* vs. *mg/+* can be subtle, all mice were phenotyped at the same age by a single person. Genomic DNA was made from a tail biopsy of each mouse and analysed for multiple simple sequence length repeat polymorphism (SSLP) markers. The first ~100 mice were typed for a series of polymorphic Mit genetic markers (Deitrich, W.F. *et al.*, 1996, *Nature 380*:149-152) from distal mouse chromosome 2 in order to accurately delimit the position of *mg*. With the first ~100 mice it was determined that *mg* mapped approximately 15cM proximal of Agouti between markers *D2Mit19* and *D2Nds3* (FIG. 13). All remaining animals were genotyped for *D2Mit19* and *D2Nds3*. Animals recombinant in that interval were typed with all available Mit markers between and for the ever growing number of markers developed during the project which, finally totaled 265 markers.

9.2. RESULTS

The murine mahogany (*mg*) gene is known to act in a dosage dependent manner within the agouti pathway, to compensate for the agouti overexpression and for lack of signaling from the nul allele *McIr* (Miller, K.A. *et al.*, 1997, *Genetics 146*:1407-1415; Dinulescu, D.M. *et al.*, *Proc. Natl. Acad. Sci.*, in press; Robbins, L.S. *et al.*, 1993, *Cell 72*:827-834). The phenotype of mice homozygous for both *mg* and a null allele of *McIr* (recessive yellow, $McIr^e$) is yellow, the same as the phenotype of $McIr^e/McIr^e$ mice, indicating that *mg* is not acting downstream of *McIr*. A similar experiment was performed with obese *Mcr4* knock out mice (FIG. 11). For both sexes, all the animals homozygous for *Mc4r-/-* were approximately equally obese and were heavier than the mice wild-type at *Mc4r* independent of the genotype for *mg*. This data strengthens and confirms the *McIr* data previously published, strongly suggesting that *mg* acts at or upstream of both melanocortin receptors.

To test whether *mg* acts specifically within the agouti pathway, experiments were performed to determine whether *mg* can suppress the obesity of other monogenic obese mutants of the mouse and whether it could suppress diet-induced obesity. Appropriate genetic crosses were set up to product mice segregating *mg* and one of the mouse obesity mutations Cpe^{fat}, *tub*, or $Lepr^{db}$ such that all combinations of homozygous and heterozygous animals were on the same mix of genetic background. No suppression of obesity was seen for any of the monogenic obese mutants (FIG. 12) lending credence to the assumed specificity of action within the agouti pathway. To ask whether *mg* can suppress diet induced obesity C3HeB/FeJ-mg^{3J} and C3H/HeJ mice were placed, at weaning, either on normal chow having a physiological fuel value (PFV) of 3.63 kcal/gm with 9% fat, or onto a high fat diet having a PFV of 4.53 kcal/gm with 42.2% fat. Food consumption and body

weight were measured weekly. Converting the grams of food consumed to calories indicated that C3H/HeJ mice on normal chow and high fat diet consumed ~97 kCal/week and ~96 kCal/week, respectively. C3HeB/FeJ-*mg*^{3J} mice on normal chow and high fat diet consumed ~83 kCal/week and ~81 kCal/week, respectively. Despite the equal calorie intake, the C3H/HeJ mice on the high fat diet readily gained more weight than the C3H/HeJ mice on normal chow ($p=0.0004$). In stark contrast, the C3HeB/FeJ-*mg*^{3J} mice on either diet showed no statistically significant difference in weight (FIG. 12D). Female data showed the same trends, although there was no statistical significance between any of the mice on either diet.

10. EXPERIMENT: MAPPING AND SEQUENCING OF THE MAHOGANY GENE

This section describes experiments wherein the murine mahogany gene was genetically and physically mapped to an approximately 0.6 cM interval, and then sequenced. The murine *mg* sequence obtained was then used to isolate and sequence the human *mg* gene. Northern and *in situ* analyses of *mg* expression in mouse tissue are also described, and sequence motifs of the predicted MG polypeptide are discussed.

10.1. MATERIALS AND METHODS

Physical Mapping: More than 36,000 individual sequences from the region were compared by BLAST (Altschul, S.F. *et al.*, 1990, *J. Mol. Biol. 215*:403-410) to publicly available sequence databases and analyzed using GRAIL (Guan, X. *et al.*, 1992, *Proc. Eighth IEEE Conference on AI Applications*:9-13) to identify potential coding sequence. In addition, sequences from overlapping BACs were assembled using phrap (Sing, C.F. *et al.*, 1998, *Genome Res. 8*:175-185; Ewing B. and Green, P., 1998, *Genome Res. 8*:186-194; Gordon, D. *et al.*,

1998, *Genome Res.* 8:195-202), and the resulting contigs were also analyzed using BLAST and GRAIL to aid in gene prediction. This data was displayed in ACEdb (Durbin, Richard and Mieg, Jean Thierry, 1991, A *C. elegans* Database, Documentation, code, and data available from anonymous FTP servers at lirmm.lirmm.fr, cele, mrc-lmb.cam.ac.uk, and ncbi.nlm.nih.gov) to further visualize predicted exons and their relationships to each other.

Northern Blot Analysis: PolyA+ RNA was extracted from the tissues indicated from wild-type, C3H/HeJ and the three mutant alleles of *mg*, C3HeB/FeJ-*mg*^{3J}, LDJ/Le-*mg*, and C3H/HeJ-*mg*^{L}, according to the manufacturer's instructions. RNA STAT-60 (Tel-Test, Inc., 1511 Sounty Rd. 129, Friendswood, TX 77546) was used to isolate total RNA. PolyA+ was isolated using Poly(A)Pure™ mRNA purification kit (Ambion, Inc., 2130 Woodward St. #200, Austin, TX 78744). 2 μg of each mRNA was separated on a 1% agarose-formaldehyde gel, transferred to nylon, and hybridized with a probe for *mg* corresponding to nt 990-1406 of the murine cDNA sequence with Rapid-hyb Buffer (Amersham LIFE SCIENCE, Gaithersberg, MD). Filters were washed with 0.11x SSC, 0.1% SDS and exposed to KODAK X-omat film overnight.

10.2. RESULTS

A positional cloning strategy was undertaken to identify the *mg* gene. Multiple genetic crosses were set up to produce second generation mice (n-2437) segregating *mg* which were used to genetically localise the *mg* locus (FIG 13B). When the genetic map critical interval for *mg* was resolved to ~0.6 cM physical mapping was initiated. Approximately 1 Mb was contiged with 30 BACs (FIG. 13C), most of which were made into random sheared libraries for shot gun sequencing. At completion of the project it was estimated that 85% sequence

coverage across the interval had been achieved and that all genes within the region had been found. Twenty-nine genes were identified, 15 of which are novel genes. Within the final minimal interval for *mg*, indicated by the arrows in FIG. 13, there were eleven genes of which nine were unknown. All of these genes were tested as candidates for *mg* by examining the three mutant alleles of the mahogany locus, the original allele, *mg*, that arose in a stock of Swiss x C3H mice, and two alleles that have independently arisen on the C3H background, C3HeB/FeJ-mg^{3J}/mg^{3J} and C3H/He-mg^{L}/mg^{L}. Each gene was examined by Northern blot analysis and RT-PCR analysis of RNA from tissues from wild-type and *mg* mutant mice, by Southern blot analysis of DNA from wild-type and *mg* mutant mice, and by SSCP analysis of genomic PCR products designed to cover the intron-exon boundaries of much of each of the genes. In all, 20 genes were analyzed in this manner, one of which showed a northern blot difference between the wild type and mutant alleles (FIG. 14).

The wild type expression pattern of this gene gives three bands of size ~9 kb, 4.5 kb, and 3.8 kb, of which the larges message is the most prominent (FIG 14). The smaller two bands can be seen in all tissues but, depending upon tissue, may require extended exposure. Each of the different *mg* alleles gave a different expression pattern. C3HeB/FeJ-mg^{3J}/mg^{3J} has extremely low expression, the 9 kb message only being very faint in brain, hypothalamus, and fat on northerns. C3H/He-mg^{L}/mg^{L} expresses a single aberrant band of approximately 9.5-10 kb in kidney, heart, muscle, fat, and, most prominently, brain and hypothalamus. The LDJ/Le-*mg*/*mg* shows an altered ratio of the three wild type messages: the 9 kb message is reduced, while the two smaller messages are more highly expressed, in particular being very abundant in fat and hypothalamus. *In situ* analysis was used to look more closely at *mg* expression in the brain and specifically the

hypothalamus. Overall hybridization in LDJ/Le-*mg/mg* looks equivalent to that of wild type, and the C3HeB/FeJ-*mg^{3J}/mg^{3J}* shows an overall reduction of expression. Close examination of the hypothalamic region in both wild type and mutant alleles revealed differences in the ventromedial hypothalamic nucleus (VMH). Both C3HeB/FeJ-*mg^{3J}/mg^{3J}* and the LDJ/Le-*mg/mg* have reduced VMH expression (FIG. 15) which is particularly interesting as many neuropeptides and receptors known to be involved in body weight regulation are expressed in the VMH, including *Mc4r*.

Initially, two overlapping mouse cDNAs of 1051 bps and 2419 bps were identified. Using these cDNAs as a starting point it was possible to build over 7990 bps of human sequences, using both the public EST database and an in house database, as well as identifying one cDNA clone from a human liver library. The 23 ESTs used in the contiging are listed in Table I below. Using the derived human sequence, it was then possible to estimate the intron-exon boundaries within the mouse genomic sequence. These were verified by PCR amplification and sequencing. In total, 4079 bps of mouse sequence was obtained, of which 4011 bp are coding sequence. The mouse genomic locus spans over 160 kb, and has 31 identified exons, at least one of which is differentially spliced.

TABLE I

Gene Bank Accession #	Clone ID #	Clone Source
NA	NA	Human Endothelial Cell (MPI)
AA062169	482948	Soares mouse P3NMF19.5
NA	NA	Human Liver (MPI)
AA350292	151062	Infant Brain
R87660	194640	Soares Fetal Liver Spleen 1 NFLS

T69367	82898	Stratagene Liver
T92696	118881	Stratagene Lung
H11351	47626	Soares Infant Brain 1 NIB
AA350293	151062	Infant Brain
AA297697	149184	Fetal Heart II
AB011120	NA	Human Male Brain
AA297214	129808	Embryo, 12 week I
AA298732	184690	T-Lymphocyte
AI076479	1676623	Soares Total Fetus Nb2HF8 9W
AA771958	1359202	Soares parathyroid tumor NbHPA
R84298	194640	Soares Fetal Liver Spleen 1NFLS
D81046	1178923	Human Fetal Brain (Tfujiwara)
AA378603	183010	Synovial Sarcoma
D60710	962349	Clontech Human Fetal Brain (#6535)
D20236	pm1235	Human Promyelocyte
AA345684	147210	Gall Bladder I
H45413	182870	Soares Breast 3NbHBst
AA044305	486349	Soares Pregnant Uterus NbHPu

The mutant mahogany alleles were also sequenced, checking all intron-exon boundaries. A 5 bp deletion at 2809 nt was found in the coding sequence of the *mg* gene from C3HeB/FeJ-mg^{3J}/mg^{3J} which introduces a stop codon a position 937, two codons 3' of the deletion. This mutation will result in a seriously truncated protein lacking many interesting domains, as discussed below. The mg^{3J} allele is the same allele that showed extremely low expression levels. The combined Northern blot analysis, *in situ* hybridization

analysis, and sequence analysis of the mutant mg^{3J} allele strongly suggest that this gene is the mouse mahogany gene.

The 4011 bp of open reading frame (ORF) of mouse MG predicts a 1336 amino acid polypeptide with molecular mass of 148,706 D (FIG. 17, top sequence). BLAST searches of the NCBI and SwissProt protein databases identified two human paralogues with a similar modular architecture (KIAA0534, Genbank accession no. 3043592; and MEGF8, Genbank accession no. AB011541), as well as a *C. elegans* homologue (YC81_CAEEL, Genbank accession no. Q19981).

Another human protein, Attractin or DPPT-L (Duke-Cohen, J.S. *et al.*, 1998, *Proc. Natl. Acad. Sci. U.S.A. 95*:11336-11341), appears to be a 1198 amino acid residue, approximately 134,000 D, secreted splice variant of the MG polypeptide. An alignment of the predicted MG (top) and Attractin (bottom) amino acid sequences is shown in FIG. 17. Attractin has not identified as being involved in the regulation of body weight. Rather, the protein is reported to mediate an interaction between T lymphocytes and monocytes that leads to the adherence and spreading of monocytes that become foci for T lymphocyte clustering (see Duke-Cohen *et al.*, *supra*).

Searching the MG polypeptide with the SMART domain tool (Schultz, J. *et al.*, 1998, *Proc. Natl. Acad. Sci. U.S.A. 95*:5857-5864) revealed sequence motifs that may provide further clues to its biological function (FIG 16B, FIG. 17). The single transmembrane spanning MG protein has a large extracellular sequence of 1289 amino acids containing three EGF domains (Nakayama, M. *et al.*, 1998, *Genomics 51*:27-34), two laminin-like EGF repeats, a CUB domain (Bork, P. and Beckmann, G., 1993, *Mol. Biol. 231*:539-545), a C-type lectin domain (Drickamer, K., 1995, *Nat. Struct. Biol. 6*:437-439; Weis W. I., and Drickamer, K., 1996, *Ann. Rev. Biochem. 65*:441-473), two plexin-like repeats (Maestrini, E. *et al.*,

1996, *Proc. Natl. Acad. Sci. U.S.A. 93*:674-678), and six consecutive kelch repeats (Bork, P. and Doolittle, R.F., 1994, *J. Mol. Biol. 236*:1277-1282). Multiple EGF domains are commonly found in Type-1 membrane proteins involved in cell adhesion and receptor-ligand interactions (Schultz, J. et al, 1998, *Proc. Natl. Acad. Sci. USA 95*:5857-5864). Laminin-EGF-like modules are found in a variety of proteoglycans such as perlecan and heparin sulphate proteoglycan. As CUB domains also frequently occur in glycosylated proteins and c-type lectins are known to be carbohydrate binders, it is likely that MG is heavily glycosylated and that carbohydrate interactions are essential for its function. Many kelch motif containing proteins have been found that, like MG, have multiple consecutive domains. Such consecutive four-stranded β-sheet Kelch motifs form a bladed beta "propeller fold" that is common in many sialidases and other enzymes (Maestrini, E. *et al., supra*). Unlike the other well recognized domains, the "plexin" repeat is less well defined. It was first recognized as a triple repeat in the *Xenopus* gene plexin that has similarity to MET (Bork, P. and Beckmann, G., 1993, *Mol. Biol. 231*:539-545). Since then, this cysteine rich repeat has been found in 6 MET gene family members, three of which signal via tyrosine kinase and three of which are hypothesized to have putative signaling function via a novel conserved cytoplasmic domain. However, it is fascinating that there is an eight amino acid stretch that is 100% conserved in the four proteins shown in FIG 16A from human, mouse, and *C. elegans*. The conservation of sequence across such widely evolutionary divergent species strongly indicates a functional domain, possible a putative signaling motif.

The multi-domain structure of MG is complex, but draws many similarities from receptor and receptor-like proteins. The full-length MG polypeptide is predicted to be a large membrane-spanning protein with multiple extracellular domains

that may have a binding or gathering function as well as a highly conserved putative signaling motif in the cytoplasmic tail.

The present invention is not to be limited in scope by the specific embodiments described herein, which are intended as single illustrations of individual aspects of the invention. Functionally equivalent methods and components are within the scope of the present invention. Indeed, various modifications of the invention, in addition to those shown and described herein, will become apparent to those skilled in the art from the foregoing description and accompanying drawings.

All publications and patent applications mentioned in the specification are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1 (FIG. 2A), SEQ ID NO: 8 (FIG. 8A), SEQ ID NO: 10 (FIG. 9), SEQ ID NO: 12 (FIG. 10), SEQ ID NO: 14 (FIG. 18A), SEQ ID NO: 16 (FIG. 19A), or SEQ ID NO: 18 (FIG. 20A).

2. The isolated nucleic acid molecule of Claim 1, wherein the nucleic acid molecule comprises the nucleotide sequence of SEQ ID NO: 1 (FIG. 2A).

3. The isolated nucleic acid molecule of Claim 1, wherein the nucleic acid molecule comprises the nucleotide sequence of SEQ ID NO: 8 (FIG. 8A).

4. The isolated nucleic acid molecule of Claim 1, wherein the nucleic acid molecule comprises the nucleotide sequence of SEQ ID NO: 10 (FIG. 9).

5. The isolated nucleic acid molecule of Claim 1, wherein the nucleic acid molecule comprises the nucleotide sequence of SEQ ID NO: 12 (FIG. 10).

6. The isolated nucleic acid molecule of Claim 1, wherein the nucleic acid molecule comprises the nucleotide sequence of SEQ ID NO: 14 (FIG. 18A).

7. The isolated nucleic acid molecule of Claim 1, wherein the nucleic acid molecule comprises the nucleotide sequence of SEQ ID NO: 16 (FIG. 19A).

8. The isolated nucleic acid molecule of Claim 1, wherein the nucleic acid molecule comprises the nucleotide sequence of SEQ ID NO: 18 (FIG. 20A).

9. A vector comprising the isolated nucleic acid molecule of any one of Claims 1-8.

10. An isolated host cell genetically engineered to express the nucleic acid of any one of Claims 1-8.

11. An isolated nucleic acid molecule comprising a nucleotide sequence that hybridizes to the complement of SEQ ID NO: 1 (FIG. 2A), SEQ ID NO: 8 (FIG. 8A), SEQ ID NO: 10 (FIG. 9), SEQ ID NO: 12 (FIG. 10), SEQ ID NO: 14 (FIG. 18A), SEQ ID NO: 16 (FIG. 19A), or SEQ ID NO: 18 (FIG. 20A) under stringent conditions comprising hybridization in 0.5 M $NaHPO_4$, 7% SDS, 1 mM EDTA at 68 °C.

12. A vector comprising the isolated nucleic acid molecule Claim 11.

13. An isolated host cell genetically engineered to express the nucleic acid of Claim 11.

14. A method of producing a mg gene product comprising culturing the genetically engineered host cell of Claim 10 so that the mg gene product is expressed in cell culture, and recovering the mg gene product from the cell culture.

15. A method of producing a mg gene product comprising culturing the genetically engineered host cell of Claim 14 so that the mg gene product is expressed in cell culture, and recovering the mg gene product from the cell culture.

16. An isolated gene product encoded by the nucleic acid molecule of any one of Claims 1-8.

17. The isolated gene product of Claim 16, wherein the gene product comprises the amino acid sequence shown in Figure 2B (SEQ. ID NO. 2), Figure 8B (SEQ. ID NO. 9), Figure 9 (SEQ. ID NO. 11), Figure 10B (SEQ. ID NO. 13), Figure 18B (SEQ. ID NO. 15), Figure 19B (SEQ. ID NO. 17), or Figure 20B (SEQ. ID NO. 19).

18. An antibody that immunospecifically binds the gene product of Claim 16.

19. A method for diagnosing a body weight disorder in a mammal, comprising: measuring the level of *mg* gene expression in a patient sample and comparing the level to that of a control sample, so that if a difference between the levels is detected, a body weight disorder is diagnosed.

20. A method for diagnosing a body weight disorder in a mammal, comprising detecting a *mg* gene mutation contained in the genome of the mammal that correlates with presence of the disorder.

21. A method for diagnosing a body weight disorder in a mammal, comprising: measuring the level of mg activity in a patient sample and comparing the level to that of a control sample, so that if a difference between the levels is detected, a body weight disorder is diagnosed.

22. A method for identifying a compound that modulates mg activity, comprising:

a. contacting a compound to a cell that expresses a *mg* gene;

b. measuring the level of mg gene expression in the cell; and

c. comparing the level obtained in (b) to mg gene expression level obtained in the absence of the compound;

such that if the level obtained in (b) differs from that obtained in the absence of the compound, a compound that modulates a mg activity is identified.

23. A method for identifying a compound that modulates a mg activity, comprising:

a. contacting a compound to a cell that contains a mg polypeptide;

b. measuring the level of mg polypeptide or activity in the cell; and

c. comparing the level obtained in (b) to the level of mg polypeptide or activity obtained in the absence of the compound;

such that if the level obtained in (b) differs from that obtained in the absence of the compound, a compound that modulates a mg activity is identified.

24. The method of Claim 22 or 23 wherein the compound identified is capable of treating a body weight disorder.

25. A pharmaceutical composition comprising the compound identified by the method of claim 24.

26. The use of the pharmaceutical composition of Claim 25 for treating a body weight disorder in a mammal.

27. The use of the antibody of claim 18 for treating a body weight disorder in a mammal.

28. The use of a *mg* antisense, ribozyme or triple helix molecule for treating a body weight disorder in a mammal.

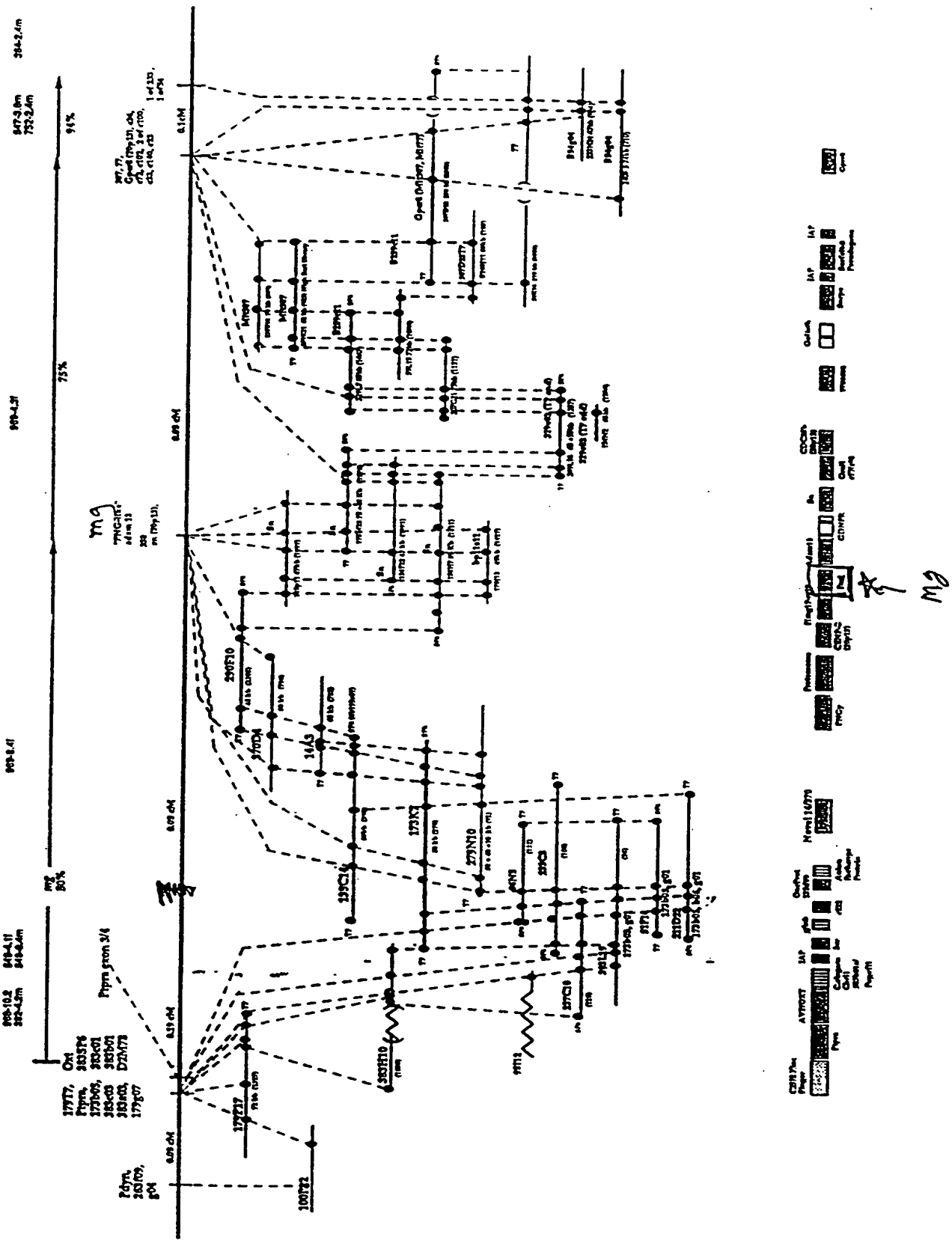


FIG. 1

GAATTCCGGGCGAAGGGGAGCCGGCGTGCGGGGTGTGTATGTGTTCGCTGGGCGCCGGCTCAGCCCCAGGAAGATGGTG
GCGGTGGCGGCGGCGGCGGCGACTGAGGCGCGGCTGAGGGGGAGCACGAGGACGACAGCAGCGCCTGCGGGCAGGAAGG
GCAGGCAGCACCGACCCTGCACCGCGACAGGGGCCTGGAGGCCGGGACCGCGCGCCCGGCTGTGTCTCCCGCGGGTGCT
GTCGCGGGCGCTGCCCCCGCCGCCGCTGCTGCCGCTGCTCTTTTCGCTGCTGCTGCTGCCGCTGCCCCGGGAGGCCGAG
GCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCG
GCCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCG
CTTCAGGACATCTGTCTCACGCCTATAATCACAGCTGTTCGGAAGGTGAGGCTGGAGGAACAGTTCGAGGCAAGCTTCG
GCTACAGAATAAGTTCAAGAGTAACCTGGGGCAACTTGGGCTTGTCTCCAAAACCAAAATGAGCGAAAAGGAGCAAGCT
AGAGTCTTTTGGGAAAATTTTAGCTGACTAATTTTTCACCGAGAACTAACTGGCTCTTCTGGATTTGTAACAGATGGAC
CTGGGAATTATAAATATAAGACGAAGTGCACATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGCTTCAA
CCATTTTGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCACCTCTGATTGCTGCC
TTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCCTGAGGTCACTGTCACTTCAGGTTATGCACTGC
TGCATTTTTTCAGTGATGCTGCTTATAATCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTG
CTCAGGCCGAGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTGGAAAGGGGAG
TCGTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGGCATCTGTAATGCAAGCGATACCAGAGGGT
GCTCCTGCTTTCCTCACTGGCAGGGTCCTGGATGTTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGA
ATATTCTGATTTAAAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGGCGGATAT
ATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGACTTCTAGGGAATGGCTTCCACTAAACCATTCTG
TGAACAGTGTGGTTGTAAGATATGGTCATTCTTTGGCATTACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGA
TTCAACAGGGAACGTGACCAATGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACTCCGAAAGCT
AAGGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGGCATCTGGCCGTGTGGTCATGTTGGTCATCT
TCGGTCATTGCCCACTCTATGGATATATAAGCGTTGTGCAGGAATATGACTTGGAAAAGAACACATGGAGTATATTACA
TACTCAGGGTGCTCTTGTGCAAGGGGGTTATGGCCACAGTAGTGTTTATGATGACAGGACCAAGGCTCTGTACGTTCAT
GGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGATGACCTCTACAGATACGATGTGGATACTCAGATGT
GGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCATACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGG
AGGGAACACACACAATGACACTTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTTCATGGCTTATGACATTGCT
TGTGACCGATGGTCAGTGCTTCCCAGACCTGAGCTCCATCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTGTACA
ACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCCTCAGTGACGTCTTGGTCTTTACCTCGGAGCAGTGCGA
TGCACACCGCAGTGAAGCTGCTTGTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGT
ACCTCCTGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAGTGTTTTTCTAAAAGAACCCTTGACCATG
ACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATACCAATGACTGCCACTGGTGCAATGATCACTGTGT
CCCTGTGAACCACAGCTGCACAGAAGGCCAGATCTCCATTGCCAAGTATGAGAGTTGCCCCAAGGATAACCCCATGTAC
TACTGCAATAAGAAAACCAGCTGCAGGAGCTGTGCCCTAGACCAGAACTGCCAGTGGGAGCCCCGGAATCAAGAGTGCA
TCGCCCTGCCGGAAAATATCTGTGGCAATGGCTGGCATTTGGTTGGAAACTCGTGTCTGAAAATCACTACTGCTAAGGA
GAATTATGACAATGCTAAATTGTCCTGTAGGAACCACAATGCCTTTTTGGCTTCCCTCACATCCCAGAAGAAGGTGGAG
TTTGTCCTTAAGCAGCTTCGATTAATGCAATCATCTCAAAGTATGTCCAAGCTCACTCTGACTCCATGGGTTGGTCTTC
GGAAGATCAATGTGTCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGA

FIG. 2A(1)

GCCCAGTGATGCTGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATCAACCCT
CTCAATGGCAGCGTCTGTGAAAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGACAGCGT
GTGGCGAGTGCACTAGCAGCAGCTCGGAGTGCATGTGGTGCAGTAACATGAAGCAGTGTGTGGACTCCAATGCCTACGT
GGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTGT
ACCTGCAGCCATTGCTTGGAGCAGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGAAAATGTATTG
AGGGCAGCTATAAAGGACCTGTGAAGATGCCGTCACAGGCCTCTGCAGGAAATGTGTATCCACAGCCCCTTCTGAACTC
CAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAGTGCAACGGACACAGCAAA
TGCATCAACCAGAGTATCTGTGAGAAGTGTGAGGACCTGACCACGGGCAAGCACTGCGAGACCTGCATATCTGGCTTCT
ATGGTGACCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAATGGGCACGCATCACTGTGCAACACCAACACCGG
CAAGTGCTTCTGTACCACCAAAGGTGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAAC
CCTCTCAAAGGAACATGCTACTATACCCTTCTCATTGACTATCAGTTCACCTTTAGCCTGTCCCAGGAAGACGACCGCT
ACTACACAGCCATCAACTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAAAA
CTTCAACCTCAACATCACCTGGGCCACCAGCTTCCCAGCCGGAACCCAGACTGGAGAAGAGGTGCCTGTTGTTTCAAAA
ACCAACATCAAGGAATACAAAGATAGCTTCTCTAATGAGAAATTTGATTTTCGCAACGATCCAAACATCACTTTCTTTG
TTTATGTCAGTAATTTCACTTGGCCCATCAAAATTCAGATTGCCTTCTCCCAGCACAGCAACTTCATGGACCTGGTACA
GTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCGCTGCTTCTGGTGGCTGCAGTGGTCTGGAAGATCAAGCAGAGCTGT
TGGGCATCCAGGCGGAGAGAGCAACTTCTTCGGGAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTCTGTAAACGTTG
CCTTGGAAACAGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCC
GTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCCTGGG
CAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAG
CCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATCTGATGCTGGGGCCAGGGACTCTCCCACGC
ACGAGCTAGTGAGTGGCACACCAGAGCCATCTGCAGGGAAGGGCGTGGCGGGGAAATGGCTGTGCGGTGCGGGACGGAA
GACTGGAAACCCTCAAAGCATCTGACTCACCTGCATGATCACAAGCTTTCTTTGACGGTTTCTCCCATCCGTGTTCCAG
CATCTAACCTTTTACTTTTGCATAGGAAATACTTGATTTAATTACAGGTCCAGGGATGAGCTGATGGTTGCTGGAGGAG
GCCAGTGTAGAGCCAGTGAGAGAACTAGGAATGACACTCAGGTTCACTGTGGAAAACTGTTCTTGGGACTGTCTCAACT
GTGCAAAAAACAAAAGATGGAGTGTTTACAAGTAGACATTCGTCATCAGTTGTTCTTGAACATGGTCTTTTAAAAACTA
GTCAGATGAATTAACTTGTTTTCATCTGAAGCCTGCTATCTTTTTTAAAAGATGTGCTATTTATTCTTGCACGATTTAG
GCAATTATCTCTCTTCCAGGGAGTACCTTTTTTTCTAGTTGAGAATTAATAATGGTCCATCTCTTTTGATCATATCAAG
CTAGGATAGAAGGGGGGCTATTTTAAATGTCAAGGTCAGCAGTGTTACTTTGAATGTAAACTGGTATAATAGGTAGTTT
TCTATAGTAACTTGATTAATTTAGTCTTAATCCATTTGAAACTCTCTCTTCCTTTCTCTCTGCCTGTCCCTCTCCTTCT
CCATCTCACCCTCCCTCTCTCACACATACACACACAAACACATACACACAACACTAAGTGCCTAGACTTTAAATAGATC
TAGCAATTGGAAAGTTAGTAAGCCTAAGTTTTTACATAATTGCATTCCTACATTCTTGTAAAATTTAAATAGCTACCAT
TGGCAATCTGCTTTTTTTCTAAAATCTGATTTGCAGCCAGGAAAGAATTTTCTCACCCAAGGAACATTTGATCTAGCAG
CAGGGATGAGAGGAAAGCAGAAATGAATGAACTGTGAAAGCTCCTGTTTTTATTATCAAAAAGGACACTGTCAAGAAGG
CGCCCCCTGCCCCCACCCCCGTGTCACCCTAGGCCTGATAAGCGATCAGAGGAAAGGACTCATTCATGTCACGCTTCCT
TGAGCAGAAAAGAGCACTGAGAGCACTTGGGACCCCTGGATCAGAGAGCATCTGTGTGTCCTGCAGCCTCCTCTGAACT
TGTGGTTCATTCTCAGGCTGGGGTGGACTCAGATGCCAGGAAAGGGACAGCCTCCCATTGTCAGGCAGAAGCTGCCCAA
AGCCTGGAGAAGGACTTGTTTGCCCTCTTTCCCCCAGGAGGGGCTCGACCCACCCACCCTCCCTCTCAGACCAAGGTGG
TGGCTGTGAGGAGGGCAGCAAATGCTGACAAGGATGAAAAGCACATGGAAAAAAATGGACGAGGAGGGAAAACTCTGCC
AAATGGAAAATGACCAAATTTAAGAGGGTGGGACAGTCCCCTGCTCCTCTCCCAGAGGGCACTGCTTGGAAATTGTGTT

FIG. 2A(2)

TTCCCCATTTATGGTGCTCTGTATTCTGGCATTATGCAGCAGCCTCCCAGAAGCTCTCTTCTGCTTCAAAACCTGGGAT
CTCTGGCATTACCCTATTGGGATGGACCGCTGGACAGCAATGCTCGAGTTTGTGAATTTGGAGAGATACTCAAAAGAGC
TAAAACTGCAGCATTTTACCTTTAAATGCAGTGCCTAGAGAGAGAGTATTGTCTCTTCCCCAACACTAACCCCACTCCC
ATGAAGAATTGCCTGGAAAGATGTTTTCAAGGAATTTGAACCATAAAACACTATCTGATGCACAGAACACCTCTACTTT
GAGACTCACCTCTCATAAAGCTTCTTTTTCACATTACTGTTAAAGACCAGACGTTCTAGAAAAGACCCCTCCTCTCATG
AGCTCCCCCATCCCTGCTACAGAACACAGCACCCATGGCGCCTGCAGTGGACTGGCCCCTTAATTCCCACAGGCCCCCC
CAGCAAGGCCAAAGGGAGGCCCCTGGGTATTGTCCTCCTACAAGGAAGATCCTCTTTGTTTGTTCAAAGGACCAGTTTT
CCTAGGCCAAAGAAGTCTCTTCCCCATGTTAGTCCTATGCCTTGAAATATCATGCACCATGACCCACAGCCATCTGGTT
ATGTCTTATTTTTTTCCTAAAAGATAATGTTTATTTTTAAAAAGGAAGGAAGAAGCAAGTGAAGTTTCATTCTGCTCCA
GCGGTGGGGAAGCCGCTGAATCCACCTGCTTCTCCTTTGCAACCGACAGCAAACAGCTTTCTCCGGCCTCAGGGCAGAA
AAAGGGAATGGCAGGGAGTAAGAGGCGCTGGGCTCGGAGCCTGTTTCCAAGAAGGAATTGGTTGTCATCTGGCAGTGTT
GCGCGTCACAAGAGAGCCTGTATATAAATTAAAATAGTCAAGACAACACTGACCTTGCACTTGTACATAACTATACAGT
AGTGTCCAGAATGTTCAGACATTCGGAGTGTACATAAAACAGAAAAAATCTTCATGTATTTTTATTAAATATAACAATG
TCTGAGTTTCACCTAAGATGTTTTTGTGCCATATGCTGGATATCCAGGTTCTCGCCAGGCCCCGATACATGAATAACAA
ACCCAAGAAACGCATCCCCATTGTGTGATGTGTTCAGATGCATCTGGCACCAATTAGGTATTTCTTAAAACAGGACTCA
TCTGTCAGAGTGCACATGAAAAATCAGGCAGGGAATCGAAACGACAGCGCTGGAGGAGACTCAGGAAGCAGAGGCGTCC
CTGCCGCTGCCCTTGGCCCTGCAAGCACATCATGACCCTTTCTGGCAGCCTCTTGGTGCTCTGGGTAGTGAGGGATGAC
CAGTCTTGTCCTGAGAAATGTTTCTCTTAGTCTTTAAGTTCAAAGACTAACCTGTAGCAATCAGACTTTCCAAAAGGGG
GTTCTCCATTTTTTGTAGTTTTGTCTAAATTTTTAATGACCATTTCCTGGAATCAGTTTATTATACTGAAAACTGGGGG
TGGGAGTAGGGAGCTAGTTTGTTGATAAATAGTTCCCATTTCCCCGTGGAGAATTTGACATACCCTGGACTCCTGTGTG
CCTCCTGCCATCCCTGCACACAGCCTGGGGAGAAGCCTGTGCCTCCCCGTGTGGAGAGAAGGCAACCCCAGATCCCCTG
AGCTAACCCGGAGGAAAGGCAGTCCTGGACAGAAGACTGTCAGCAGAAGGAAAGTACTGGACTACCCGTGGGTAAGTCC
TGCCATTCAAGACTGGAGACACCTGGGAAATAAAAAGAGCAGGGCACTGCTGGTGGGAAGAGGCATTTTACCTTCCAGT
GCAAATCCTGCTCCTTTGATTTAATGGGGTGTACTGGGGCCAGGGGCTGATTCACTTCCTTGGGAGATGGTGGTGTTTT
CATGAACATCTTTGATCCTTCCATTTCATTTATTCATCCATCCATTCAACAAGTATTTGCTAAACACTAACTTAAGCTA
ATGCTAGGGTAGTGACTGAGATGTAAAAATAGATTTTAGAATTAAAACAAAATCCAAGTCCTCACACCCCTGTCATCCC
AGGAGATCTTTCCTTGTGGTGGTTTCTGTGAGAATTGGCCATCCTGAGGACACAGCCAGGACGGCAGAGGCCTCCTGGC
CTCAGGGCATGCCCTGCCTACCTTCTGAAATGTTTACCCCATTGACCAAACTTGGCTCCAGCCATTGCGGTGGTTTCTA
GATAGCCAGGCCCACCAAGAGATATTGCCCCTTGATGAGAGTCAAACACCCTGCCTACAAGGAGATGTTTTGAAATGGA
GAGGAAAATTGGCACCTCATCTTTTAAAGGCAGTAATGGAATTGATTTTCAGTAACTGAATTTGTGCACAAAACATTCT
AAACACTAGTGAAGCCTGTTTCGTTGAACTAATTCTGGCTCTGGAAATGTTTTTGTTTTATAGTTATTTACGATTTCGT
TTGTTTGGATTCAAGCTTAGTTTGTTAATATGTATAATTTAGCATCTATTACACTCATGTAAATATGGAGTAAGTATTG
TAAACTATTTCATTGCGGGGATTGTGGGTGTTATACATACATTTAGGACTGCAATTTTTTGGTATTTTTTGTATTGTAA
AATAACAGCTAATTTAAGCAGGAACAAGAGAACTAAGGGAGGTCTGTGCATTTTAAACACAAATGTGAAGAACTTGTAT
ATAAACAAAAGTAAATACTATAATACAAACTTCCTTCTGAAATAAAAGTAGATCTGGT

FIG. 2A(3)

MRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNETAPEVTVTSGYALLHFFSDAAYNLTGFNITYNFD
MCPNNCSGRGECKSSNSSSAVECECSENWKGESCDIPHCTDNCGFPHRGICNASDTRGCSCFPHWQGPGCSIPVPANQS
FWTREEYSDLKLPRASHKAVVNGNIMWVVGGYMFNHSDYSMVLAYDLTSREWLPLNHSVNSVVVRYGHSLALHKDKIYM
YGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKDQYAVVGHSAHIVTLASGRVVMLVIFGHCPLYGYISVVQEYDLEKN
TWSILHTQGALVQGGYGHSSVYDDRTKALYVHGGYKAFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSG
TMLVFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPELHHDVNRFGHSAVLYNSTMYVFGGFNSLLLSDVLVF
TSEQCDAHRSEAACVAAGPGIRCLWDTQSSRCTSWELATEEQAEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHW
CNDHCVPVNHSCTEGQISIAKYESCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGNGWHLVGNSCLK
ITTAKENYDNAKLSCRNHNAFLASLTSQKKVEFVLKQLRLMQSSQSMSKLTLTPWVGLRKINVSYWCWEDMSPFTNSLL
QWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNGSVCERPANHSAKQCRTPCALRTACGECTSSSSECMWCSNMKQCV
DSNAYVASFPFGQCMEWYTMSSCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQASAGNVYP
QPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCEDLTTGKHCETCISGFYGDPTNGGKCQPCKCNGHASL
CNTNTGKCFCTTKGVKGDECQLCEVENRYQGNPLKGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMF
INASKNFNLNITWATSFPAGTQTGEEVPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPIKIQIAFSQHSN
FMDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQLLREMQQMASRPFASVNVALETDEEPPDLIGGSIKTVPK
PIALEPCFGNKAAVLSVFVRLPRGLGGIPPPGQSGLAVASALVDISQQMPIVYKEKSGAVRNRKQQPPAQPGTCI

FIG. 2B

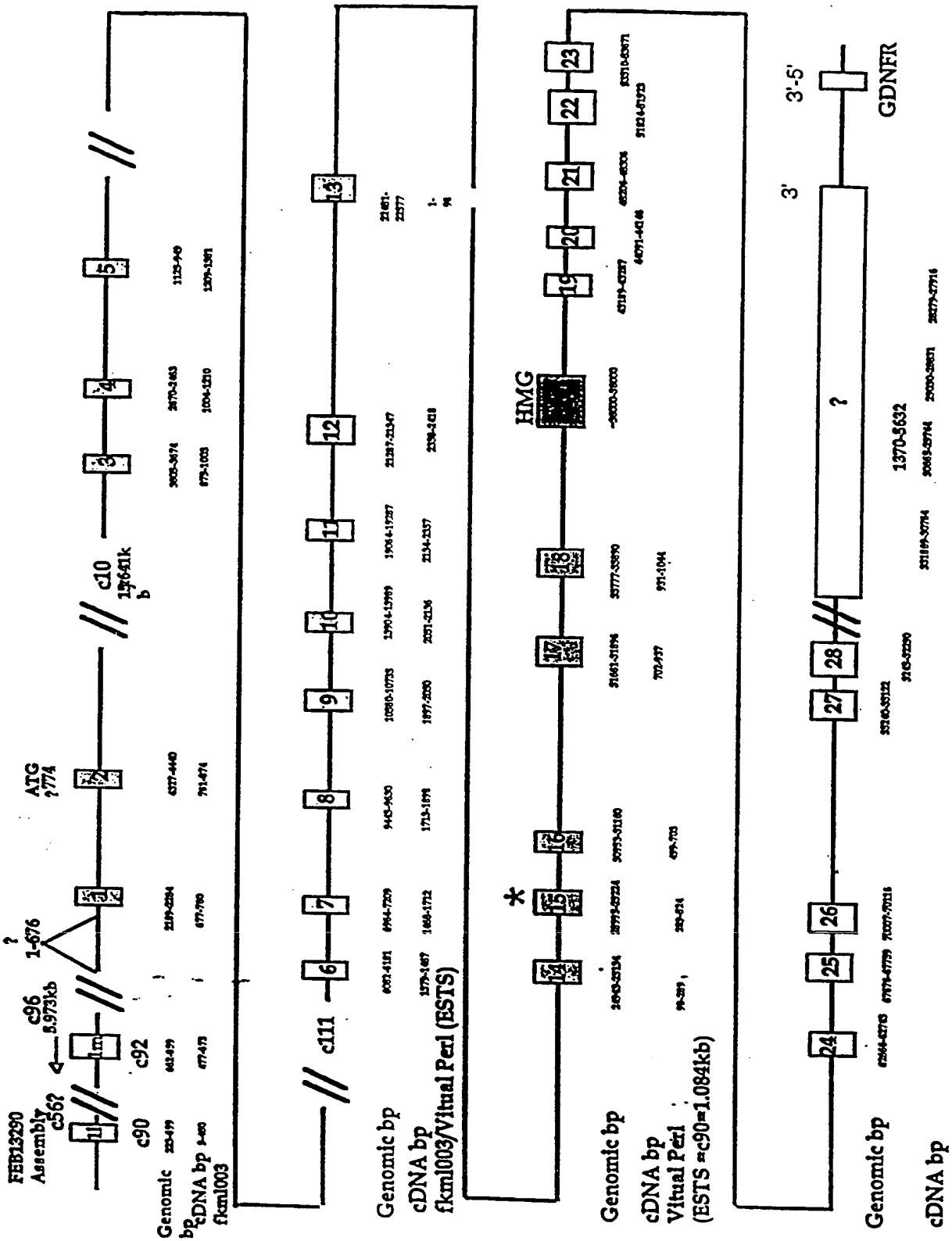


FIG. 3A

AGATTTTATGCCTTCGTACACGCCTCCCATAAGATGGACAAGGTGTACTA
ATTACTGCCATTACTGTTGCTGACCCCAGAGGTCAATGTCCTCACATGGC
CTCTACTGGCACTGTCTGGGCAGAAACTGTATATCCAACTGGTGAACCTG
AAAGCCCTATGACTACTTGGTGTCTCTGGTGCTAACCCTAGTCGTTGGGG
CATCTTACTGTATCCTGGTAAGGAAAGACATCCAGGCTCCCCACTTAYMK
WWACYRGYWMRGMYCAKGSYMGRGCYAAWKTKCTGTRRMRTCTGGCTGGC
ATAGAGACATTACTATTGAAAGTTTTGTCTTTCTAAATCCTTGGACTAAA
GAGAGCACAAGATTTTCTGGAAGATCTTGCTTTAAATTTTTTTTTTATTC
TTTTGAGATGCTACATATAATTAGAGGCCCTGCACATGGAGGCGAGAACC
CCACCTCTGGGCTACATCCTACGTCTTTTCCTTAGGGTATTTTTTTTTTCT
TTCTTGTACCTATCAGTATTACTAAGTTGCAAATGTGCTCAGCAGTAAAT
TTAACATACATAGGCAAAAAGAAAAGTCTCAGGACACCCTGCCTCACACT
GTTTACTGTGCTCAGGAGTACTGAGCCATACTGTTTTCTTGCTGCTGCTT
TTTTTCTCTTGGTTGTTTACACACAGTGTTCAAGGTGTGTTAATCATAGT
TAGTATTTCAATTTTTTCTTAGGTCAGCAAGAAAGCTCACAGAGGAAGAG
TGCTTTGCTGCCAGCCTGATGACCTGGGTGACCCAAGTGATCTCACCTAC
AGGGTGGGAGCACAGCACAGCATTCCAAGTCTTTTTCTGACCACACAGGC
ACTATGGCACACAAACACACAGGATACATAAATGTTAAAAAAAAAAAAAG
ACTTTTATATTTTTCTCCATATAATTTAAAAGATTCCTCTTTCAACATTC
CTTTTGCAAAGCAGTATCATTGTGTTTGTATATGTGTGTCCTTCCACATT
TTGTCTTCAATTCTAAATTTTTAGAATTGTTAGCCTGGTCCTCTCATTTC
TACTACTTTCTCTAGTAAACTGTCCTTTCATATTACACATCGCTCTCCTG
TCACCTGTTTTAGAGCTGTCATCCATTTTATAAGGTTACTTCACTGTTCT
ACACTACTTTGTGTCTTTTAATTACTATGCCTGGGGTGATTCAAAAACTG
TCTGTGATGGGTTGGTTGAGGATGGCTCCAATAGGTTCAACACTTGGTCC
TGATTGGTGGAACTGTTTTGGGAAGGATTAGGAGGTGTGACCTTGTGGGG
GAGTGTGTCACTGGGAGTGAGTGACCTTTGAGGTTTCAAAAGCCCATGCT
AGGCCCAGTGTCTGTCTGCCTGTCTGTCTGTCTCCTCCCTTTGCTCTTTC
TTCCCTCCCACTTGCTTGCAGATCAGATTCGAGCTCTTAGCTACTGCTCC
CGTGCTGTGCCTTGCTGCTACCATGCTTCTTGCCATGATGTTCATAGACT
TACTCTCTGAAACTGTAAATAAGCCCCCTAATAAAATGCTTTCTTTTAAA
ACTGCCTTGATCATGGTGTCTCTTCAAAGAAATAGAACATTAACAAAAAC
ACTATACCAAACTGCCTAATAGTCCTACTAATTTTATGATGAGTGCTAGT
GCTTTATAATCACTAGAAGAAAAAATTTCCAGGCCATAAAATTAACATGG
TTTTAAGTATGTATAAATCTTGTCTTGAAATCTGTTTTCTATAACTAACT
CTAATATGATAATGTATATTCTACCTTCAAAAAAGCACAAATAAGACTTC

FIG. 3B(1)

AAACCCTGGGAATTGTTAGACAAAGGCCATTTAATACTAATAAGCTATAA
ACTGAAACCATCTGATATATGAAAACTATTAATAAAATCAAGATAAAATA
ACCCCTATTTATATAACTTACTATATACCTAAAGCAAAATATCAAAGAAA
GTACCTTAAAAAGATAAATTATTCTTATTTTGACAATGAATTCTTTGGGG
CGTTAAATTGTAGAATATCAACACATATCAAGAAAGTTTAGAAGAAAACT
ACCAAAGTTTAAACAGACTTTCCTCGGTAATTACTGGTGATTTCTTGGCT
TTTTTTTTTTACACTGCAGTTTTTCAGGGTGGAAACTTAAGCTTTGTACA
GAAGCACTTACCACCACTCTCAGAGCTGGAAATGGCTCAAAGGGCAAAGC
ATTACAAGCCTGGCAACCTGAACCAAATACCCAAAACACTTGCAAAGGTG
AAAGGAGAAAACTAACTCCAGGAAGTTGTCCTTCGAGCTCCTCTTGCACA
CCACTGTATACACCCCCTTATATACACTCAGTTACCATAAATAAAATGTT
TCATTATAAAGACACTTACGCTAAAACCATGCTGTAATCTGAATGGTTGA
ACATATATCCGCCAACAACCCACATTATATTTCCATTGACCACAGCTTTA
TGAGAGGCTCTGGGAAGCTTTAAATCAGAATATTCTTCTCGAGTCCAAAA
AGACTGGTTAGCTGGCACAGGAATTGAGCATCCAGGACCTAATAAAAAAA
AAAAAAACAACAACAACAACAAATAGCTTCACAAAATGCAGCCTGAAAGT
TTATAGTATTCCAAGTTCCAATCTAAGTGCAAAGAATATTTAAAGACTTG
TGGGGCTAGAGAGATGGCTCAGTGGTTAAGAAAACTGACTGCTCTTCTTG
GAGGTCCTGAGTTCAAATCCCAGCAACTACATGGTGGCTCACAACCATAT
GTAATGGGGATCTGATGCCCTCTTCTGGTGTGTCTGAAGACAGCAACAAT
GTACTCACATGAAATAAATAAATTAATTTTTTAAAAAACAGACCAGAAAA
AAAAAAAAAAAAAAGACTTGTGTTTCCTTTAGCACTTAAGCGCAAACATC
TTTAACTTGTGGGGTTTTAAAGGTTTTTACATGTACAGGTATTTTGTTTA
CATGTATGCCTATATACCACTTGCTTGCTTGGTACCCAATGATGTCAGGA
AAAGGCATTGAATCCCCTGGAACTAGAGTTACAGATCTTATGAGCTACTT
TGTGGATGCTAGGATCAAACCTGAGTCCTCTGGAAGAGCAACCAGTACTC
TTAACCAAGAAGCCATCTGCTTAGCACCTAACATGAGTTTTTAACTTACT
CAAGATACAGACCAAAACCAATCACTCCCTTATAAAATTTAATACTACAC
ACTTTCTGATAATTTGGCAATTTCTGATAATCAGGTTAAACTTTTTTAGA
GGTAAAAATCTTGCTGAAGCAACATTTAGTAGAAAGGGTAGACCAAGGGG
TTATTATATTAACTCATGTGGAAAAGGCATTAGGGTTGAAATATAATGAC
AGATCAAAATCGATCTTCTGGCAAGTCCAGGCGCTGAATAGATGAAAGAG
ACAAAGGGAGAATTGGACAAACTAAAAACATTTACATGAACACTTACTTT
CTGAGGACCTAAGCATAGAAGGAAAATCACTAAACCAACGATGACTGCTT
CCTCAATACCCCAGGGAATTCCCTACAGTACCTTAGTACCCGGTTGTGTT
GGGTAATGGCACTAGATGACAGCACTGAGACTCTAAGGAACGCTTGTCCT
CCTCTCAGCTTGAGTCTCTGCTTCTCTATCACCAGACCATGTTCCCTAAT
TCCCACGAATGAGTTGCAAAGGATTTGTCAAACCTTTCCACAATTCTAAG
CACATAGATAACAACCACATATATGTAAATTCAAAGAATCTGAATAAATG
GAGATGAATGCTTAAATGCCACCTGATACATGATTAACATAAGGCGTATG

FIG. 3B(2)

GCTGCTAAAATAAACTCCCTACAGTTCACTAACTCAGAACTTTCTGTGAG
GGAAAGGACTTTGAAGGGCAGCTCCTACCCTGCCAGTGAGGAAAGCAGGA
GCACCCTCTGGTATCGCTTGCATTACAGATGCCTCGGTGAGGAAAGCCAC
AGTTGTCTGTACAGTGAGGAATGTCACACGACTCCCCTTTCCAGTTTTCA
GAACATTCACACTCAACAGCGCTGCTGCTGTTACTGCTCTTACACTCTCC
TCGGCCTGAGCAATTATTCGGACACATGTCAAAACTACAAAGACAGGAGA
AAACGAAGTCAACAATTTCAACTAAGCAACATTGCAACTAATGCAGACCT
TCCTCCTTCAGTTTAAGTTCAGTTCATTTGCAAGTGTGACTGCAGGACTT
ACCAGTTAGCCCAAGTGTGCTCACAGAGCTCTGTGTAGCTAGAGCCCCAG
GCTCAAGTAATGAAATCAAATCAACCTTGCTGCATTCACATATGAAGAAG
GAAGAATAAATAACTCACAAAGTTAGAGAAATTACAAAACAATAGACATT
TGTGCAAAATCACTTAGACTTAGCTCAAGACTGGCAACCAGGATCCTACT
CTTTCTGGTAGCTCATTAGTAAAGAGTTCTACAAAAGCAGCAAGGTCATG
CTAGGAAGTGGAGGAAGGAGAGGAAGCCAATGAGCTGCCAACATTCACGG
TATACATTTCTCTGTAAAGATTCTGAGAATTAACAGAATTTAAGATTATT
TTCCAGTGATGTAGTTAAAGGTCTTTAGTAACTTTTATCAGCTTAGAAGG
AGAAGAGCAGTTAACTTCATGTATGAGTTTAAGTGTCTCATGACTTAAGA
TAACAGTTTTGCTACAATTTGAAATGCCATACTTCAGACTTTTTAAAGGG
GTGCATTAGTGGACTATTACAATAGCTTAAAAATATAGATTTCTCCTACT
GATGATTATTACTGAGACACTACTAGTCTTTATTAAATTCACTTAGCAAA
ACTCCTGACATTTTCTTCCAGCAGCGGAAGAATGTCTCTCTCTTCTAGGA
GATCCTCAGTGACAAGATCTAGAAAGACCAAGAACTGTGGTCCCAACCAG
TGGGGCTGATATTTGTTTAACCTTTTAGCTCCTGTTTCTTCAATTATGAA
AAAAAAAAAAAAGAAGAAGAAGAAAATCCATGTTAAAATTTAGCAAGGAG
CCTGACTAGCTAGAAGCCTCCCTCCAATATATTAGTGTTATTAAGTCATT
TGAGTAGTATCACAAATATTAAATCTAAATATCTTACTTGTAAGTGATAT
TAAATCCAGTCAGATTATAAGCAGCATCACTGAAAAAATGCAGCAGTGCA
TAACCTGAAGTGACAGTGACCTCAGGAGCCGTCTCATTGCCATCTCTTTC
AGGAACAATGAGGCCACTGAAATGTAAACACAGACCAGATTACAGCAACT
TCAACAGAAACTGTCTATATGTTACTATTTGATCCTGCTGCTCCTGTTCC
AACACACACTGTAAATGTGACTCTAGCTGGCCTCAAATTCACAGACCCAC
CTGCTTCCACCTCCTGGGTTATAGGCATGCGCTACTATGCCCAACATCTA
AAAGGATTTGAAATCTATGACTTTGATTGAATTTTTGGTTTTTTGTTTTT
GCTATAAACTTTTTATTATAATACTCTCAAGTCTCTACAATAACATTATT
AACAAACTTTATGAATTGACAACTGTCAAATATATACTGTTGAAAGAAAA
TACTTTACATATTTTTGTAATATGTATCATATAATCTTTTTAATGTATTT
TATAGATGTCTTATATAAGTAAAAATAGAAAAGTTTACTGATTTATAATC
CTTATACTATTAGCTTTCAGACGTATTTTTGTTGTTAAACTGGTAACACA
TTTTATGTTTATAATTCACAATAAGCACTGCCACTGAAGGTGCCAAAGGC
TCCCTAGAATCTCAGTAAGAACCTAGTGGGTAATATTTGAAGTTTTGGAT

FIG. 3B(3)

GCCAGTAAATTCATGTGTAAAGATTTATTGAGTAAGTGACTACCAGCGGG
ACAGTGGTGGTGCACGCCTTTAGTCCCAGCACTTGGGAGGCAGAGGCAGG
CGAATTTCTGAGTTCGAGGCCAGCCTGGTCTACAGAGTGAGTTCCCAGGA
TAACCAGGGCTACACAGAGAAACCCTGTCACCCTGTCTCAAAAAAAAAAA
AAAAAAAAAAAAGAATATACCATTTTTAAGGCATTTGATCCACAAAATCA
TACCACCTTGTTTTACAAAAGATATATATTAACTTGAAGGCTGGAAATGG
TGGCACATGTCTTTAGTCCCAGTATTGGGAAGACAGACCCAGATGGATCT
CTGAGTTCAAGACCAGCATGGTCTACATAGTGAATTCCATGTAAGTTTGT
CCGTGTGTGTAACTTGAAACCTCATTATAGAATGGAAGTGTCTACCCCAC
CCCACTTACCAACAGTAAGGAATATTATGTTGGTCCCGCTCATTTAATAC
ATGGTGTACTCCCAAGGTAAATCATTTTCATGTTTAGTCGCTCCTATTAT
TTTTTCCATTATCAATTCACTACAACTACTACCACCAATCACATTTAGCC
ACTAGAAAAGCCATGTGATTTGCTCCACACATACAACTTCACTCAATAAA
TAAACATCTTATCAGTACTACTCTCTCTTTCACTCACTCAATCCCTAGTC
CCCTAAGTTTTTGGACGATTACACCAGGTAAATTCCTACTTCAGGGTTAT
GACCATCTTAAAAACTACGACCTAGCAATTCTCTTTGTATAAGAAATACT
TCCCCGTATATACACAGAAAAACAAAGAACACTACTACAGCACTATTCAG
ATGACAACTGACTAAAAGTCACCTAATTGCTTATTTATGGGAGTTGATTA
AATTAGTCATTACAAATCTGTAGGTCTGCAAGACTAACCAAGAGCTTCGT
GAGGACAATAGGTAGGGCTACCCAGAGAAACCCTGTCACCCTGTCTCGAA
AAAAAAAAAAAAAAAGGGAGGCACAGAGAAAAAACAACAGGCCCGGGGTA
CCTGTACATCTATGTAAGCGTAGGTACATGCACATAAAAGTGACTACAAG
AGAACATAAACAGAGAGCGCCGATGAGAAGAGGATGGGATTTTTCATTTA
ATTTGCGTGTATGAGAGCACCTATATGTGCATGTTATCCGCACCAAAGTG
TGTAGGGTACATTATGTGAGTGTGCCTGCAGACGTCACTGTCAGGTGTCT
TCAATCACTCCCCTCCTTTTTCTTCTGGAGATAAGAGTTTCATGAAGTAG
TACTGGCTGGACTAGAACTCACTATGCAAACCAGGCTGGCCTTGAATTCT
CAGAGAGCCTCTTGAGTGCTGGAATTATATGCATGTGCCGCAACACAGCC
CACCTCATTTTGGGGGGTAGGATCTTTCACTGAACCTGAGCTCACTGATT
GGTTAGACCGGACTGGCCAGTAAGTTCCAGGACCTCTCTTGTCTCCGCCT
CTTCAGCACTGTGATCACAGGCTCACAACCACACCTGGACTTTTACTTGA
GTCCTGGAGATCTAAACTCAGCTCTCCATGCCTGTGCAGAAGGAATTAAA
CTGAGCCAGCTGTCTCAGTATCAAGAGAGAACATAGGAACTGTAAGATTC
TGACAGTACTCTAGGGCTTACAGAACACCGACACATTTTCTACTATGTAT
TCAGTTAATAAAAGAATAAATACAAACAAAAAAACATGAGAAACATATAG
AGGCAGAGACAGACAGACACACACACACACACACACACACGCACACAC
ACACACACACACACACGCACTTAGACGGGTGTGGGGGAAGAAAGAGCAAG
GCCACCTAGAAACAGGTACGTTCCATGCAAATGATCACAGGAAAGGATTG
GGGATTTTTAACCACTTGTGGGAAATGCTGTACTCTCCTATTCTAGCACA
GATTTGAGGAAAAAGTAGACCAGAGAGTCTGTCCTTCCACATATCCTGGA

FIG 3B (4)

AAGTCACTGACATGTCCAAGTTTTGATTTCTTCATAGGGACAATGAGAGA
AACCCAGACTATCTCACAGCAGCACAGCAAGGACCAACCAGCAGAGCAGG
AGAAGTGCTTACAGCAGTGTGCTGCTAGAAGGTGCAACAGTCTTCTTACA
GAGGGCATTTAAATATGCAGGATGGATAAGTTTGCCAACTACAACTACAG
AGGCTGGACAAGGTAGGACAGCTTCTTCACTGTCAAAGACGTTTGGGCAG
TTGCTTCTATTTACCTTAAAATCAAACTGTGACAGCTGTGGCATATATAG
ATTTCTCCCAGAATGAAAACACATTAACTCACTTATGTCAATAATATGGA
GTAAACACAAACATAGTCTATCTAGCTCAGCATGCAAGACATGTGAGGAA
GAGGAGCTACTGTGAGTCCCTATCCCTGTCCCTAAGGAAACCAATATATG
TAAATGTAGTCTAAGCTGCAGGCAGTTCTTCAACTGCCTACCCCAGGCTG
CTCACCACTTCACATTCTAAGCACAGACTAGAAAGTATGATCAACCTCTG
AACACTGTGCTATAATGTTACCATCAATCTCACACACAAATTTCATAACA
TTTTAAGTAAGTCTATGATGATTCTATGTTGTGTCCCAGTTATATAAGAT
CCATAGGTCACAGGGTAGACATTCAAGGACACCAACATTTGGAATTTTGG
GTTTTTTTGGTGTACTGTATATACTTGCTAGTGCAGGTACCCATGCTCAT
GTGTGTAGAAGTTGGGCGTCTTTCTTCTATCACTGTCTACTTTATATTTT
CTTTATTGTTTCATTTGATATGTATAGGTGTTTTGCCTGCATATATGTGT
ATGTTTGTTGCCAGAAGAGGGTATTGAATTCCCTGGGACTAGAGTTACAG
GTGGTTGTGAGGCACCATTATGGGTACTGGGACTCAATCCTGGGTTCTCT
GGAAGGGCAGCCAGTACTTTTAATCACTGAGCCATCTCTTTAGCTTCCTT
CGTTCATTCGTTCGTTCATTCCTTCATTCCTTCATTCCTTCATTCAGAGG
ATTGAGATACCTTCCTCAGTTAGGCTGGCTAGCCAATGGACTCTGGGAAT
CTATCTGTTCAGCTATTCTCTCCTTCCCCATCCAAGTGCTGGGGATACAG
GCAGGTCCTACTGGGTTCATTTTGAAAAATTACAGAACTATGTATTTTCT
TCATAAATCTGAAACTCAGCATAACTGTCTCAGGCTAACATGGAATCCCT
AAATATATATGAGGCACAACCTGACTTTACCAACTGTACTATGTAAATTT
GCTAGTATATTAGTCAACACTTAATGGAAAAACATCTGATAAAAACAACT
TACAGGCCAATAGGCAAGGAGACACTTGGGGAGGTGGATTCAAGGCAGTC
ACTGGATTCTTGAATTTAAGTCCAGCCTAGGCTACATGAGATTCTGCTTC
AAAAAATAAACAATTAAATTTATGGGGGAAAGAATGATGTATTTTGGTTT
CAGAAATTCCATCCTATCATCCAAGGGAGATATTGTATAACAGCGAAGTT
CCTCAGCTCACAGCAGTCAGTAGCATATAGACAATCCTGGCTCCAAGCCT
ATGAAAACACAGCCTGTACTAAAGGTGTGTTCCTGTGTTTTGAGTGAGAT
GTGCCCCCTAAGTCTTGTGTATTTGAATACTTGGCACTCACTTGGTGGCG
ATTTGGGAGGAATTAGGAGGTGTGGCCTTGGTGGAAAAGGAGCATCACTA
GGGTCAAGGTTTCAAAATCCTCCTGCCATCATCCCCAATATGTCCTCTCT
GCCTCCTGCTTGCAGTTCAAGCTATGAGCTCTTAGTTACTACTTCCACCA
CCTACCCCTGCTATCTCTGCTCCATCATCATGGACTCCTATTCTGGTGGA
ACTGTTAGTCCAAAAAAGTCCTTTCTTCTACAACTTGATTTGATGCCAGA
TCTAGCCCCCCAGCCTAGCTAGCAATATACCAAGGTATACCATCTTGAAC

FIG. 3B(5)

TCTAGGTGTCTCTCAATCCAATCAAGCTACATAAGATTAACCATCATACC
TAGTCATCCCCAAATCAGTGTATCTCTCTCCTCCCAAGACTATAAGCTCC
TCAAGGGTCAAAATATGTAGAAAGGAAGAAAGATTCTCAAAGGTCAAGGA
TCAGACCTTGGTGAGGATTGAGCACTGTCTACACTTTGCCTGGTAAAGAA
GGGTCCACAATGTAAAAGAGAACTGACCTGAACAGTTTTCAATTAGGTGC
TAACAAATGTCTCATACGTATTGAGTTTCTTATAAATAAATAAATAAATA
AATAAATAAATAAGCAAGCAAGCAAGCAAGCACTTAAGAGCACTAGCTGC
TTTCTTCCTGAAGACCTGGTTTCAATTACCCAGCACTTATACAGAGGCTC
ATACCAATTGTAACTCCAGTTTGATGATATCCAACATCTTCTTCTAGCCT
TCAGACACCAAGCACCAAGCATGTAATGGTATAACACATGTATACCAAAC
ACCCATACAAACCAATTTTTAAAAAAATATTCGAGCCGGCGTGGTGGCGC
ACGCCTTTAATCCCAGCACTCGGGAGACAGAGGCAGGTGGATTTCTGAGT
TCGAGGCCAGCGTGGTCTACAGAGTGAGTTCCAGGACAGCCAGGGCTGCA
CAGAGAAACCCTGTCTCGAAAAACCAAAAAAAAAAAAAAAAAAAAAAAT
AGTCATTTTAGGGCTGGAGAGATGGCTCAGGGGTTAAGAGCACTGACTGT
TCTTCCAGAGGTCCTTAGTTCAATATCCAGCAACCACATGGTGGCTCACA
GCCATTTGTAATGGGGATCCAATATCCCATTCTGGTGTGTCTGAAGACAG
CTATAGTGTAAATAAAATAAAGAAATCATATAAATAAAATAAATAAATCT
TTTTAAAAATATTAATTAACCCAGGCTGAACCTAAACTTACAAACTTCCC
ACATTAGGCTCTTTAATGCGGGTGTTATAGGTCTGAATACCAGCTTAAGA
ATAATATTCTTCTGAAGAATGTGCCCTGGTCAATCACCATGACCACACCT
GCCAACAGGTCCTTCATAAAATACTTGGTATATGTTGAATGTTCCATAAA
ATTATGGAGCTAGAAAAGGTAGTGAGCTAGAAGGATATTAAAGATATAAA
CCATTGCCCCAGTGGTCCTCACATTTGTCTAGTAATAGAACGTTGTTAAA
CTGTTTTTATTTAGAATTTCAATATATAAAAGACAAATATGAAATAGTCC
GGAAGCAAATTAAGCTACAGCTTGCAGCAAAGCCAGATAGAATGCAGATT
AAACTAACACAGTACCTTTGTCTTATGTTTTAGATGCTAAAGTCTAGTCT
ACAACCCCAGCTGCCCTTGAACTCTTAGCAGTCCTCTTGCTTCAGCCTCT
CATGCTGCTAGGGTTAAAAGTATGTGCGACCACACACAGTTTTGAAGTTT
AGAGCACTTAAATGATCTATTCAGCAACTCAGGCAGGATTTACACTGAAA
GTAAATTATCTTATGAATCCTTTTTGGTTTTCCTTTTATTCATTTCATTC
ATGCACCTTACATGAACTATCTATTGCTAGGCTGTCTCTATACTGGATGC
TCAGCACATCACCAACATGCCGATTCTTCTACTGGTACAATGGCAATGCT
GAGAAAACCACACAACCTAAGACAGTAGGGAGGTGGTGCTCTGATTGTTG
GTGTTGTTGTTGTTGTTGTTTTGGTTTTTCGAGACAGGGTTTCTCTGTGT
AGCCCTGGCTGTCCTAGAACTCACTCTGTAGACCGGGCTGGCCTCAAACT
CAGAAATCCGCCTGCCTCTGCCTCCCAAGTGCTGGGATTAAAGGCGTGTG
CCACCACGCCGGGCTCTGGTGCTCTGATTTTTAAATACAACAATTTTCAG
CTAGCAATGTAACTCAGTAGTAAAATGCCTGCCCAGCATGCACAAGGCTC
CAGACTGGACCCTGAGCACCACAACACTTTTTAAAAGATGTGTTTATTTT

FIG. 3B(6)

ATTTTATGTGCATGAGTGTTTTGCTTACATGAATGTCTGCACTGTGTTTA
CCTGGTGCCTGTGAAGGTTAGAAGGCAATGGAGCTATGGAGAGTTGTAAA
CTACCATGTGGAAATGGAGCTATGGAGAGTTGTAAACTACCATGTGGGTA
CTAGGAATTGAATCAGGGCACTCCTCTGCAAGAACAACAAAGGCTCTTAA
CAGCTAAAATATTACTACAAACCCACACCACAAAATTTTAAATTGATAGA
CATTATCACCTTAGTTCTAGATAGAGAATGTGCTTGGCATTGTAAGTACT
AAAAAGGTTTTGGGGTGGATCTTTTATATTATCTCACTATAATTTTATAA
AATTAATACTCAAATATGTTATAAGTTAAGGTTTTTATTTTTGTTTTTCA
TTTCTGTATTTTGTCTATGTAGCTCTGCCTGGCCTGAAACTCATGGGAAC
TTGACTGGCCTCAAACTCAGAGAGACCTGAACGGCCCTGCCTCCAAAGAG
CTGGGACTAACCATGCCCAACAGTAGGTAGCTTTAATACCTAACCAGTGT
ATTAGTTCATGCTCTCAATTAACCAACATTCTCTACATACAGAAATTTTT
ATGCCTATTTAATCAAATACACAGTCTAAGTAAACTCTAAGTACAACTGC
TTGGCTCATATTCTTACAATGGCTATGGCTAGCTAATTCAAAGGCCAGTC
ACATAAAAGGGTCTCTATGAATTCTGATTAACAAATGCAGTTAAATAGAT
GAATTCCTAAAAAGTAGTATCATAATAATATCATATTTAGTTTTTGTGCT
TCCATTATAGTTTGAGGTGCCTCCTCCCATAATGCAAGGTATATTTCAAA
TAATAGATATATACATGGTTAACACATGGCAAATGCCATTTTAAATGCTT
AGCACAGCCTGCTCTTTGGCTCCATTAAGTGAAACTCTTAAGTTCTCAGT
TAAAATAATTGTTGGAGAGCTATAGGAGCAATGGGTGGAGAACTAGTCTT
CTAATTTGTCCTTTGCCTCCTTGCGTACTAAGTAGTCCCTCCCTCACTAT
GTGGCATTCCAGCAGACTACCACCAAGAGAAGAACAGAAAAGTGTTGATT
TCTTTCTAAAGTAAAGAAATAAGGGGCCAGTGAGATACCTCAGCAGGTCA
AAGCCATTTGCCTAGAAACCAAAGTTCAATCCTTGGAAGCCCTGTAAAGG
TGGAATTAGAAAACAGACTCCACAAAACTGTCCTCTAACCTCCACTCGGG
CACACATGTGCCAACCCCTCCATTCTCCCTCCCCCACATACAAAGTAACA
ATAAACTTTCAGAAAATTTAAGTTGCTACGCATGGTGATTGATGAATGTC
TTTAATTCTAGCTCTTGGGAAGCAGAAGTGGGTGGATCTCTGTCAGTTCA
AGACCAACCTGGTCTATATAGTGTGTTCCAGGCATCCAGGACTACACACA
CACACACAAAATTACGTGAAGGAAGTAGAATGTTTGAAGGAAAGAAGTCT
GGAAATGGGGATGGAGAGAGACCTCAGCAATTAAGAAAAGGTCTTGCACC
GGACGTGGTGGTGCATGCCTTTAATCCCAGCACTCGGGAGGCAGAGGCAG
GCGGATTTCTGAGTTCGAGGCCAGCCTGGTCTACAAAGTGAGTTCCAGGA
CAGCCAGGGCTACACAGAGAAACCCAGTCTCGAAAAAACCAAAACCAAAA
ACAGAAAACCAGTATGATAGGTCAGGCAATTGGATCGAGACAGGACACTC
AAGATAGCTAGCCTGTGCAATATAGAAAGAAGTCTCATGGAAGAGAGAGG
GAAGGGAAGGAGGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA
GAGAGAGAGAGAGAATGAGAGCGAGAGAGCGAGCGCACCTCAGTTGATAC
AAGATTGGGGCCCTGAGTTCCATCCCCAGCATCCCATAAATTGGGTGTAG
CAGCACACACCTGTATCCCAGCAGAGAGGCAAAAGACAAGTTCAAAGTCC

FIG. 38(7)

TATATGGAAAAAGTGTGAGATCAGCCTGGAGACCTGGTGTGTGGCAGTGG
GGTGAGGGGTGTCATCAAGGAGAAGGCTTAGTAAGTAAAGGACCTGCGTT
GGTTCTTGAGTTCAAGTCTCCAGCAATCAGAGAAAGCCAGAACCATTGCA
CAAACTTGTAAGCCAAGTGTTGGACTGGACAGAGACAGGCAAATGTTTGA
GGTCCAGGTTCAGTAAGAGACCCTATCTCAAAAAATCTGATGGAGAGTAA
CACTGGAAGAACTCAGAGTGAGTCACACATGCACACACAGGTGAATGTGT
ATACAAAGGGGGCAGGGAGGGAGAATGAGAGGAGACTGGGAGATATCTGT
AGTTCATGTCTGTAATTCTAGCACTTCAGAGGCAGCTGGAGCTACACAGC
AAGACCCCGTCTCAAAAACAAACCCAAGCCTGACAGTGGTGAGGTACACC
TTTAAGCCCAGAGGCAGGAGAATCTCTGAGTTCAAGGGCAGCCTGAGTGA
GTTCCAGGACAACCAGGGCTCCACAAAGAAACACTGTCTTGAAAAAAACC
AAAACCAACCAAACAAAAAAGAATCAAAAACAACCACCACCACTACAACA
AAGCAAACAAGGGAGAAGGTATAAAATGCTTAGGAGAGTCTTCCTTTAGT
CTCCATCCTTTGGGTACTCCTTCCCCACAGAAAGCCACTACTACCAATTT
CTTACATAAGCTGCTGTTTTAGACACAGGTTTTTTTTTTTTTTAAATATA
GTAACATATTCATGTGTAGCTCATTTTTCTAGTGAGTGGTTGGTCCTTCT
TTTAACAGTTTAAAGGACCTCTATGTTTAAAGGCGATTGGCCCTTGTCTG
GAGTATGGGTTGTATTTTCCCAATTTGTGAGTTTTACCCAACCTATTGCC
TATTACCTATGGCCATTTATTCTTGTCGATAAGTAGTTTCCAATTGTATG
ACTATGGTCACAGTGTTCCATGGACTCTTCTGCCGCTAGACAGCCCCTGG
GTCTGAATTTGAGATGGTTACAAGGGTGATTGGCTCTGCTCCCTGGGTGC
TGGGATTAAAGGCGTGCACCTCCACACCCAATTTGTTCTGTTTTGTAAGA
AATGAGGTTTTATTGTGTTGCTCAGGCTGATCTCAGTCTCCTGGCCTCAA
GGTATCCTCCCATGTCGATACACAGCACAAGGCGTAGGAAAAGTGGCAGA
TTTTTTTAAATTAAGTTTTCTTTCCAAAATATAGATTCAGAAATGTGAGA
TTTTCACAAAGTGAACCTGCTCACTTCCCTGGCTCTMGAATCTCCATTGT
GGCTCCCGCCCATCCCTTTTGCCCACCAGTGGCTGTTGTATTGACTTCTA
TCCCATTCCTTAACTATACCTGTCCTTGGTCTTCGCTGTGAACTTGCTTG
GGCTGAGAATCACCTTGTTCCGGGCACATCAGGTCAGTGAGGGTGTTTCC
AGAGAGTTTTAACAGAGACCAGAAGACCCACTCCAAATGTGGGTGGCAAT
ACCTGATGTTCTGTCATCCTGGACTGGGTAGGAAGAGGAAAGTAAGAAGC
AAACGGCACCCCCACCTCTCTGTCTGCTTCCTCGCCGACACAAAGTGACC
AGGGCCTCCCACTCCTGCCCCCTCAGCTAGAGACACTTGCTGCCATCTTT
CCAACCACTCTGAGACTGTGCCTACTAACCGTGACCCAAAATAAATGTTT
CCTTCCTTAAGGTTGCCTTTGTTAGCTCCTTTAATAGAGCGGTAGGACAT
GTAACTGCCACAGGCAGCCATCGCTGCCAGCCCCTCCCACTGACCGTCTG
AGAACCACACTCAGCTGTAGGCACAGCTCTCATAGCTGTGTGGGCGTAGC
TCTGTCTACTCGGTCATTCCCCTGCTGCCGAGCATTTATTGTTTTCAGTT
CCTGGCTGATGGGTAGCACTGTTATGAACATCCTAGTACAAATCTCAGGG
TGACACGCGCCTTCATTTTTCCTGAGAAAATGCCCAAGGATAAAATGCTA

FIG. 3B (8)

GGGCCAAGGGAAGAATATTTCACCATTAAGAGACACTGGTCAGGACTGGA
AAGATGGCCCAGTGGTTAAGAGCACTGACTACTCTTCCAGAGGTCCTGAG
TTCAATTCTCAGCAACCACATGGTGGCTCACAACCATCTGTAATGGGATC
CAATGTCCTCTTCTGGTGTGTCTGAAGACAGTGACAGTGTACCTACATAC
ATGAAATAAATAAATAAATATCTGAGAGAGACAGACAGACAGACACTGGC
TAGTCATCTCACAATGTTCTCATGTTTAAAATATGATACCATTTGTATAA
AGCAGAAACACAGGAAAAATAAAATCTGTGGTATTATATTTGATTTTTAA
ATTAACTTGATTAGTGAAGTTAGCAGCTACACTGGGCAGGGGTTGGGAGT
GGGGTACTCTGAAGTGCTGGTATTTCTGGTTTTGTTTTTTGTTTGTTTGT
TTTTTTATCTTATTTATATTACATAGAAAGCCATTTTGCTAATACACTTA
CCATGTGTATATATTGTGCTTGAATTACAGCTAAGTAATTATTTCTGAGG
GGCTTTAGACTACTGAAGATTGGGCCCAATGAGCCCCACCCCAAGTAGTC
TCCAACATCCCTCTTGGAAGTACTTGAGAGCAAAGATTCAAGTCACATGT
CCCCAAACCCTCAGCAGCCACCACCCTTTAGGTGTGGCTTTTGCTCTCGG
TCATCCTGGAACATCTTGCCATCTTTGGTTTGTTCTCTCCCTGTCTTGCC
TCTGGTAGAGCTGGGTTTCTGTGCTTCTATTCAACCATGTACAAGAACCA
TGTGCCACCTGCCATGTGCCAAGCCTGTGCCAGTCCCTGTGAGCGAGCAG
CCCACCCCGTGAGTTATCATGTGAGGAGCTATGAGGAGCAGGAAGGGGCC
CGGATGACTTCAGCAGACAGTATGAAGCAAGCACTGTGCGATTTATGCTC
CCTGGCCACATGCCCACAGATGGTGTCTGAGACACTAGCGTTTAATATTT
GAATTCTCCACATTCTAGCCTAGACATTTTGGTTGCAAGAAGAAAATTGA
CTCCAGTTGTATCCTGGAATGAAATTTATTGGAGGAAAATACTGGACAGG
CTCCCAGAGAAAATACGATATTCAGGCACAAAAAGAAATGGGGACTGAGG
ATCTGAAGTTCAAGGTCATCTGTAATGAGATTGAAGTCAGTTTGGGCTAC
ATGGGACCTGGTCTAGGGGGAATGGGGAAGAGAAGGGAAGGGATCGAGAT
AGGGAT

FIG. 3B(9)

CAATGTGCTCTGACGATTAATGGGCTAGAAATGTGTGGCTGTTGATTAGT
GAAAAGATGTCATGGTTCAGGAGATTGGTAGTCTCTGTGGGAAGACAACT
CACTGAAAGGGAGGAAATAGCCTGGAAGAGATAAAGAGACAGTGATCAGC
TAGGAAGCTTAAAATTTAAATTTTGTTGGAAGTACTGTTAGGAATACTAG
CAGAGGCCAGATGAATGTATGGTTAAGTTATAGCAAAGGAAAAGATTGTT
AATGGTGAGGTTAGGAATGCAGGGTGACACCAACCTGTAATGTCAGCATT
AGCGAGATAGAAGCAGGTGTTTAAGGCCATTCTCTGCTACTTAGCAAGTT
GAGGCCAATCTGGACCACATGAGACCTTTTTTCAAAAATAAATCTCCTTA
AACAAAAGAGGCTGGGTTTTTTGATAGATTCTTCAAGATGTTAATGTAAA
TAAATGGAAGACCAAGGATGGCATGCTAATATCCTCAGTGTCTGAAGAAG
GACTATGTAGTGTTGGCTGCTGACTCTGAAGTAAGTGCTCATTACTGACA
GATAGTGTATCTTAGAGCCTGGCAGATGGGATGGAAGTGAGGAAGCAAGT
AGCACCTTTGTATATTATGTTCTAAGTAGCCAGAGATACTTGACACAAAA
CAAAGTTGAGAAAATGTATCTTCTAGAAAATACAGACATGGAAAGGTGTC
CTTTCTATAAAAGAGGTATTAAACATTAACCTGAAAAAAAAGTTAGCAAA
TTGGGCTTTGGCAAATGAATATAGTCAAGTTTCATTTTTATTTTGTTTTT
TGTATATGACTGTTTGGCTTGTTGTACCATGTGTGTTCCTGGTGCCTAGG
AAGTCAGCTGGAGTTACAGATGGTGTGAGTTGCCATGTGGGTGCTGAGAG
ATGAACCTAGGTCCTCTGGAAGAGCAGTTAGTGCTCTTAACCACTGAGCC
ATCTCTCTAGTTCCTTCTGTAGAATTTTCATTAATTTACAAAGGAGAAAG
TATAAATGATAAAACCATGAGAAGATAGACCGGCACTAGAATTAGTGGAG
TCAAAATGTTAATGATATGTCAGATACGCCTTATATGAGGAAGTTGCAAA
ATTATGAAAATCCAGGCACTCCACTGAGTTAGAAATCTAGGCTCTGATGC
ATACTGCTATGGTAAGGTAGCAAGTGGCCATTGAGTGCAGAAGTGAGTCT
GGATGGGTCTTCTGGTGTTGTGGAGCACACAGACTGCTGTCTTCTGCATT
GCAGTTTCACCTGTATTTCCTTGGAACTACTTAGCTTTGCAACTAGGCGT
TAAAAAAAACTTTATATTTATGGTTTTAAGTTATTTATTTGTTTTATTTT
ATTTTATGAGACATAGTCTCACTCTCTAACCTAGGCTGGCCTGGAACTGC
CTAGGTAACTTGAGCTGGTGATTCTCTTGCCATAGCCTTCTAAAATTTTA
GATTGCAGGCATAAGCCAGACCACTCCTGACTTTTGTAGCCATTTTTCTG
ACATGAAGTGTAACTTTGCTTTCATAACTAAAATGATTTAGTTGTTTTGT
TATTGTTTAATCCCTTTTGCTTTGAATGTATCCTTTTGTGTGGGTGGCAG
ATATATAACCACAGACTTTTCCACAGGCATCCTACCCTAGGTCCAGAAAT
GACTCTGAGACGTCTTATATATGAATGAATGCCTAGGCCAATAGCTTTGG
CTGATTTCCACGGGTTCATAGCTCAGTTATCCCATTTAAACTAGTCTAAG
TCATGCCATGAGGCTACATACCCCTCCTTCAGTTTCAGGCGACTGTCTTC
TCAGTTGTGTAATGTCCTATCCTCTGTTGCTGCTCCCCAACCCCCATCCT

FIG. 3C(1)

TGCGTCATAGTCCGTCTGTCCTCGTCTCCCCCCATTTACTTGCACAACGG
ACTCTACTCTAGAAGTCCTCTCTGTGCTGGAGCTTGCACCTCCGCTCTCC
CCGTCTAAGCTAATAGGCAACAGCATTGTACAGACAGGTGATGCTTCCAT
ACATCGCACAGGAGATTCTCCCTACACAGATACTTATTCATCCAGCGTGA
ATGCAACCGTCCAGGCGTGTTCTCCTAGTTGTAGTACATGCTGTTGTATC
AGTCTGATGAATTTCTTTGTCTTTACAACCAAGAAAGATAATACTGTAAG
AAATTTTGACTAACATTTTTTCTTTATTTAAATTACAGACTAACTGGCTC
TTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACGAAGT
GCACATGGCTCATTGAAGGACAGTAAGTTATAATGGCTGACTTTATTTTA
ATTTATTATAAGAGCACAGTATAGCACAAAATACTTCCATGTGTGTTATT
GCTATTTCTTGAGACAGGACCTTTCTGACTGAGTAACTCAGGCTGACCTT
GAATTTTGCTATGCTACCTCTGCTTCCCAAGTGCTAGGGTGGTAGGTGTG
GACCACCATGCCCTGCTGCTAAAATACCGTTCATTGATGCTTTTCATTTG
GATAGTGTTCTTGCTTTTTAAAATTTACTTTTTGGGGGACCAGAGAGATG
GCTCAGTGGGTAAAGTGCTTGCTGAACAAGTCTGGTTATGTGAGTTAATC
CCTGGCTCCCACAGTGGAAGAGTGACTCCTGAAAGTTGTCTTCTGACTCC
CACGCTTGTGCATGCACGCACACACACAAATAATAAAATAAAAAATTAAA
AGGAAATTTTCTTTTTTGGGTGATAGGGATTGAACCTATGACTTCACTAA
GCAAGTGCTCTATTGTTAAATAATTCCTTTAATTTGTGGGTTTTTTTTTT
TTAGGTTCCAAGTTGACTTAATGTTATAAATGAAAGATACATACCAGAAA
TTTGCATATTTCTAATAGTTTAAAAAACTTAGTTAAAATCTTTTTAATAG
TTTGCTTAAATCTTTATATAATAATGCTATTATATCATTTTTCTAAATAT
TGATTTTATTATCAGCAAAACAGTAAATGAGCCATCAGAATAACCACTGT
AGCCTGTTTCCCTGGCCCTCTGTCCTTCCATCTGTCTATCTTCTCTTTTT
TTTCCTTTTTTGTGCCTGTCATTTAGGGCAAAGCATTTTAGTCTCTGAAC
AAAACTTTGAAATTTCCAAGTAACTCTTGTTTATTTGTTGTGTCTCATAT
TCAACCCAAGAAATATTATTTACTAACTCATTTAAAAGCAACAATTATAA
CCCACTACATGTTAGCAGAAAAACCTATTTGTTTTTATTGAGACGGGATC
ACACTAGTAAGCACTACATGGCATGGCGTTCACTGTGTAGATCAGGCAGG
CTGGCTTCGTGCTCTTGACAGTCCTCCTGTGTTTGTCTCTCACTTCTGAG
TGCTGGGATTATAGACATTACCAACACACCGATTTGGGGGGTTGGGGTAC
TGGGATCAGTCCAGAGTTGCATGGATGCTAGGCAAGCACTCCACCAACTT
AGCTATATCCCTGGTCATAAATGTCATAAGGAAAAAAATTCCTTATATTT
AAAGAAATTTTAAGAATTGCATTGTTTAAGATTTCACAGATCTCTTTGCT
ATCTGGCAATCTTTTTTGATATTTTGTTTTGTTTTTAAAAATATGTGGTA
TGTAAACAAACTTAAATATGAATGGGACAGTTCCAGATGAGAGTGAAAAG
TTAAATATTTGGGAGAAAAATTGATAGGTTTATCTATTATGGAAAATTTC
AGAGATTTTAGTAAAATTTGAAAATGGAGCTGGGAGGTCTGAGGTAGTCA
TCTAAAGCTGCCAGTTGTAGAGCGTGTTGGAGTGTGGAGTCAGAGGGAGT
TACTGATACACTTGTTGAAATTGCCCAGGCTTCATGGGAAGTGATGAGGG

FIG. 3C(2)

GCTGTTACTGTGACTCTGGGCAGGGCTTGTTAGTTTCCTTTGGATTTAGT
CTCAGTCAGAGTTGATACATAGTTTCCTGAGGACGTGGCTTTTTGGTACA
GTGCTGTGAAAAGGCAGAGAAGCAGGTAAACTTAGAAAATGTGTGTTTTT
AAAGTGATGTGTTATGAAATCTTACGTAAGATGAATAAAGAAAGAAGTGG
GGACACTGAGGGCTCCTGTTTCTAAATGTTAAAAGCAAGGCTGGAAACAT
TCTTTGAAGGCCCCTGAAGTCAGAGCCCGTGTCTCTTTTGGTTCCCAGGA
CATTTTTGATATTCCCTTACACATAGCAAATACTAACTAGATCTCTGACA
AATGCAGGAAAGCTGTTTATATTTATATATATTTATATTGTATATTTTTC
TCCTTATAAATTCTTTAAAAGTCTGTTTTAGTAGTTAATGTTATGATTAT
TATAAATTACTTAATTATTTTTCTAGGCCAAATAGAATAATGAGACTTCG
CTTCAACCATTTTGCTACAGAATGTAGCTGGGACCATTTATATGTTTATG
ATGGGGACTCAATCTACGCACCTCTGATTGCTGCCTTTAGGTAAGCCCTG
CTGCATTTCATCTCAGGAAGTAAGTGTGTCTCCAGGATGGAGTCCGTGCT
GCATTTACTTTATTCTGCAGTCACACTCATCTCATGGAATTAGTTCTGTT
CTGGTGAGCTACAGTTCACTTGGTTTTTATGTACTGGGTCGTTTTCCATG
TATACTAGTATGTAGCCACGGTTAGTCTTGAACTTCTGGTTCTCCTGCCT
CCACCTTCCAAGTGCTAGGAGTATAGGCTTGTGCCACTGTGCCTGACTCA
TTTCACATTCTTGAACTGTGAAGTTTTGATAACACTATTAAATTTACCTG
CTATTTGTGATTTTGTTAAAGTTTGCATTAAAAAGTTTTTGACTATATTG
ATAATATTTTTGTGACAAATTTAAATCAGAAACCATACCTTTCTTGTTCT
TGTATGTATTTCATTCCATAGGCCCTTAGGAATAACTTTTTTCAATAGTA
TATAGTTCTCTCAGTTTGTATATATGTATTATTAGGGATAGGAGGAGCTT
TCTGGAAGACTATTTATAAATTGGACAATGGCTAGCTGTTGAGAGTGAGG
AATTTGCTAGTTTTGTTTTGTAAATCCCTCCCCAATGCATCTGTATTAGT
GATTTAATAAAATAATGCAATTTTGTCAGTTATATGGGTTGCACTGAATT
TTTGCTATTTTATTTTAAGAAAGATTTTTGTGTGTCTACAGTGTATATGA
GTGTATGATATGTGTGCGTGTGCATGTGTGTGTGTACTTCTATGCAGGTA
CTCACATGCTATGGTGTGCACGTAAGGTTGGAGTGCAGCCTCACATGTTG
ATCATTATATTCCACCTTGTTAGAGATAGGTTGTCTTTGTTGTTTGCTGC
GGCCTGGAGCTGGAGCTGGAGCTAACGAGTCTCAGCCACCTGACATGGGT
ACTGGGAACCAAGAGCAGCAAGACCTCTTCTTCTTCTTCTTCTTTTTTCA
TTTTCGGTTTTTCAAGACAGGGTTTCTCTGTATAGCCCTGGCTGTCCTGG
AACTCACTCTGTAGACCAGGCTGGCCTCGAACTCAGAAATCTGCCTGCCT
CTGCCTCCCAAATGCTGGGATTAAAGGTGTGTGCCACCACCACCCAGCCT
AAAAGATTTTCTTACTAAAATATATTTCTAAATTAATTAGTTGGAATCTG
GTTCATACTTCTTTTTGAAACAAAACCAGCATTTTTTTTCATTCTACATA
CAGAGACATTGACACTAGACACTGGTTATGAGTAGTTACTATAAGAATGG
GAAATTATTCCACCCTTGTAAAACTTAATACAACTCCTTATCAGGCTCTG
AAGACTTTTTAAAAGCAAGAATTGTATATAACACACAGAAATGATTTAGA
CTATTTAGATCTTTATTGCATGGGATTTTAAAATTATTATTGTATTTCGT

FIG. 3C(3)

GGGCATGTTTTGTCTATGTAGCATATGTGCCTGTAGAGGCAACCACCAAG
TAGGTCCTGGGAATCAAACCTGGTACCCCTGCTCTTAGGTGTTCTTAACT
GCTGAGCCATCTCTCCAGTCCTC

FIG. 3C(4)

AGGCAAGAAAGAGCCAGCGAGCCTCCAGACAGACCATTAGAAATTCCACA
GTCAGCACAATAGGGAGAACAGTAAATCTTACATTAAAAGAAGGCCAGGG
CCTGGTAGCAAAAGGTTTTAATTTAAGCACTTGAGAGGGAGAGGAGGCAA
ATCTCTCTGATTGGGGGTTGGGGTTAATGGTGAATGCCATGACACCCTGC
TCAGAGTTAGCCTTCTCCCCTAAAAAAATTTTAAATTCATTTTCAATGCT
GACACAGTTAATCATAGACATTGTATCTCAGACACCTCAACATACTCCAG
ACTGCAGCACCAGCCCACTGCTGAGGCTGTCGTTCAGTTGGTAGAAGGCA
TGCTCAGCATTCGCGAAGCACCAGACTTCATCCTTAGCACTACATAAAAC
TGGGTGTGGTCATGCACACTTATAACTTCAGCACCATGGAGGCAGAGGCA
GGATGATGAGAACTTGAGGATCATTCTCAGTTACATAGGGAGTTTGAGGT
TAAGCAGGGGTACAGGAGGCCTGTCTCAAACAAACAGACAAACAGACAAA
CAAACAAACTTCAAAAAACTCTTGAAGTACTAGGCCTAGTACGTGCTGAG
ATTGTAGGTATATGTCATCATGCCTGTTGTAGAATGAGTGAGAGCGGACT
CCATAGGCTTATAGATTTGAATCTTGGTGTCTGTCTATGTCATGTCATCC
CTGCACAAAAGCCCACACTAGGCCCACACATTCTCTCTGTCTGCTGCATG
TGGGTTAGATGTGAGCTCTCAGCTGCTGCTCCAGTGCCATGCCTGCCTGC
TGCCAGGCTCCAGCCATGACGGTCAGGGACTAACTCTCTGAAACTGTAAC
CAAGTTCCCAATGAAATGCTTTCTTTCATAAGTTGCCTTGGTCATGGTGT
CTGCTTCACAGCAATAACACGGTGACTAAGATACCTGGCTCCTCCCCTCC
CCACCCCACCATTATTTACCATAAAGTAAACAATACACAGTTGGATAACA
TGATACTGAAGTTATTTTCCTGTTTCCTGATGTAACCCAATTTTGGACAA
GATTAAGCCTTAAATAGCAAGCTGTGAGGCAGGATAAAGAAAAAGCTCGC
AGGCCAATGTCTGCTTTACCAAATTCTGTTCAGCAGTCTAAAGCTGCCGT
CACCTCGACTCCTGTGATGGCATTTCCATCACTATCTTAGATATTCCCTG
GGTCACAACCTTTTAGTACACAGATTGCAACTCTGATGGAATGGCTGACT
GCTTGGCTAATTAAAGCAAGCTAGAGTTTGTCTGGCTTCCTTGTCTGAAT
GGGGAGGTGGTATTTACAAAATTTTGTAAATAAACTACTATATTTGCATG
ATGTATATAAATTTGATGTGGCTGCTTTTAAATCATTTAACCTAAACTGT
CCCACAGAATCATCTGTTTGATTGGAAAGATTGTAGCTTCAAGAGAATTT
CTGCTGAACCTGAAATGATTCATAATGATGTGTCTGAAGAATGTGTGCTA
TCACCTACGGTTTTTGTTTTAGTTGATATTTGTACTTTAAGATTTCCTTT
ATGTATGTGTGTGTGTGTATTTATGTGAATGTATACCTCATGTATGTG
GTGCTCAAGGACACCTGAAGAAGGGCTCTGGAGCTGGAGTTACAGGGAGT
TGTGAGTGCTAGGAAAGAAAGCTGGGTACACTGGGAAATCAAAAGGTGCT
TCTAACCACTGAGAAATCCTGCCAGCCCCTTGGTTTATTAAAAATATCAA
ACAAAACCAACACTAGTTACATAAGTATCTCTCTCTCTTTCTTTCTCTCT
TTCTTTCTCTCTCTCTTTCTCTCTCTCTCTCTCTGTCACACACACACACA
CACACACACACACACACACACACACAAAGGATCCATAATAGTTCTTCTGT
ATCCCGGTTAAATATAAGTTCTTAGGGGCTAGAGAGATGGCTCAGCAGTT
AAGAGTGCTTGTTGTTCTTCCAGAGGACCCAAGTTCAGATCCTAGTACAC
ACATCAGGCAGCTCACAGCTACCCATATCTCCAGCTCCAGGAAGAACCAA
TCAATGCCTATGGCCCATGCAAGCACCAGCACACATATGCTCCACAAACA
TCCATATATATAGCTAAAAGTAATAAAAATAAATCTTCAAAAAATTAATT
CTGGTTGAACTGAAAAAGATCACCTAACATTTAGAAAAAGCAGTTTACTA
GTGAATAGGACATAAATCATGGTATCAAATATTCTGTTGTTAAAGGAAGC
AACTAGAAAAAGCATGTGTTTGAAATAACCAATGGATACAAAACAAATGA

FIG. 3D(1)

GGCAACCCCAACATCTGTCAGTACCTTGCAAACCAACACAATAAATTTGA
TTTTATTTAAATCGTAGTTATTTTTCATGCTAGTAGTTTTGAAACACAAT
AAATTTGATTTTATTTAAATCGTAGTTATTTTTCATGCTAGTAGTTTTGA
AACCAAGATCTAGATTTTGTATAGCCACATAAATACACATTAGAATTGCA
AACTGATACGAGCTTCATCTTCATCAGTCTCTCTTCATGAAAAGCAGTTA
CAGGGACTGAGACATGACTCAGCAGTTACGGCATGGGCTGTTCTTCCATA
CGACATGGATTCAATTCTCAGTGCCCAAATGTTGGCTCACAACCATTTGT
AACTCTGGTCCCAGGGGATCTGACACTCTTCTTGGCTTCTATGGCCACTG
TATTCATACGGTACACAGACACATATGCAGGCAAAACTCAACAAAAAAAA
TAAGGTTTAAAAAAAAGAATTAGAACTTAAAGGCACTTCATTCCGTCAGC
ACTAAATCAGCCTCTCTGGAGTCTTCCCACTTCATGAGAAAATCGTCAGC
TCTCCACTGCTGTCTGTGGCTGAGGAGCAGGACCTGGACAACGTTCAGAG
ATTGTCAGTGCATCTCTTTTCTTCTTTGGTTTGCTGTCATCAGGTTCACT
GTCACATTCCCTTTGTACCATCCTTCCTTTAACAGCCTTTTGAAAATGCA
GAAATGTTGGATGCTGCCTTCAGTTCACACAGGCTGTCTTTTTAGCTCCT
CATCTATCTATGCTTAATTTGTTAGTGGTGCTCACCCATGTATGTGTTTA
TGTCATGAAGCCACAAGATGAGCCTTGATTGAGTCTTGCTGTCAGTGTGG
ATCACAGAAATGACACCCTATCATCTTTGCTTCCTGCTTGTTAGAAGTCA
TTGATTCTGCTTATACTCAAGGCCCACAGTATTATACTTGGGTGTGAACC
CCAGGAAGCAGGGAGGTGGGGGGTGTCATGGATACTACTCAGATATCTGA
CTGTTGTGATATTTCATCAGTTCTCATTGGTCCTATCTTTAAAATCTGCC
CTACATCTAGAGCTGGCTGTGGTGGTGTGTGTGGTGGCATCAGTATCAGA
ACTTGGATTACAGAGGCAGGAAGATTGTGATTTTTGAGGCCAGAATAGGT
GCATACAAAGATCCTGTCTGCAAAAGAAACAAATGTGCAAATAATTATAA
CTACTTTACTAATAGCCTAACTAATAACCACTGCTAGTGCTGTGTCCACG
AAAAGGTGAAGTAAACTGTGAAAATGACTTCCCCTTCTGTGTGACACACG
CCGTCATGTGATTTTACTTGTGTCTCATCATTGTTTTTCCTTCTGTTTGC
ATGTGTGAATGTTCACATGTGGAAGCCAGAAGTCAGTGTTGAGTGTCTTC
ATAATTGATCTCTATTCTCTTTGTTTTGAGACAGGGTTTTGAGACTAAGC
CCAGTGCTCAGTGATTCATCCAGTAAACTGTAGGGAGCTTCCTGTCTCTG
CCTCCACAGTGTTGGGATTACAAGCATGATCCAAATTATGTGACAAGCGC
TTTACTAACTTAGCCATGTCCTCAGCTCCCCACTCCCCTTTTCTTTTCTT
CTTCTTTTTTTTAGACTTACTTGTTTATTTTTATGAATGTCTTGCCTGCA
TGCATACACACACACACACACACACACACACACACACACCCCACATGC
AAGCAATTCCAGAAGAGGGCATTGAATCCCTGAAACTGGAGTTCCAGTTA
ACTGTGAGCCTGTCATGTGCGTACTGGGAGCTAAATCCGGGTTCTCTGGA
AGGTCAGCAAGGTCTTACCTGGGAGCCGTCTCTTTAGCTCATGTGTTTCT
CTCTTGAAGCAAGAAACCTAGGAATCATTTTGAAACTTCCTTCACAGCCT
TTATCATAACTTCACGTCAATTTTTACCTACTCTTTCAACAAATACATGT
TATATTTACTTATTTTTATGTTTAGCCTGCTATTGGTTTCTACTTAGCCT
CTTGCAGTAGAGTTCTGTCAGATTTATGTTTCTATTGCTTTTAATTTATT
TGTAAAGGTGAATGGGAAAATATTTAAAAATTACAGATCCCATCATTTAC
TATATTCTTAAAAGCCATGGCTAGCCAGGCTTGGTTGTGCATGCTTGTAC
TCCCAGGACTCTGACAACTCAGTAAGGAGGAGAGTGAATCAGAAAATAGC
GCCAGCCTGTGCTGCTTAGCAAGAAACAGAAACAAGTACAATCACACACA
TAGAAAATCCCCCATTAATACCATCCCATTAGATATAATGGTCCTGTATG
ACCATTCAACCACTGTTTGTCCTCTGTACTGCAGTAACAGTCTTCTGCCC
TTGCCCGTGAAGCACGTGCGCACCCCGCCTCCAAGTGCTTTTGCACTGGT
GTCCTCCGTCTAGATGTCCTGTTACTATATGTAAGGACTGGTTTCTCCTC
CTCTTTACAGTTCAATCTAATTGTCTCATGAAAAGATCTTTCCTGACCAT

FIG. 3D(2)

CTGGTTCAGACAGGTTCTCCCTGTTGTTGTTTTGTTTTTTGTTTTATAGT
TCTAAATTCCTTTCAGGAACTTTTGCTTATTTTAAATTCCCTGAGTGCAT
ACGTGTGCTTGTTGTTGCTCATGCTCGTTGTTTGGGCTTACTTTACTATC
AGCTCTGGATGTGGTTCACAGAAGGTGCTCAGGGGAGCACTCTCAGCCAC
TCATCTCACACGGGTTATAGATATATGTATTGATGCTACGTTTGCTTGTG
AGCCATGTTTTAAAGATTAGAATATCTTTTCTATGTGTACTCTATCAAAA
CACATGTTAGGGCTTTATCTATTTTATACAGATATTGGTGTTCTTGCTTT
ACTAATTTTCATGGAATTTCGGTGAATATTAGTATTTTAGATAGGAAGAC
TTGTCTCAAAATGTAGCTCAGCTGGTTGAGTGCCTGCCTGCATGTAGAAA
GCCCTGTATTCACTCTCCAGCACCTCAGAAGTGGGCCATGGTGCATATGC
TGTCATCTCAGCACTCCGGAGGGAGAGAAAAGAGAATCTGGAGTTCAAGG
TTATCCTTGGCTATATAACAAGTCCAAGATCAGCCTGGGCTACATGGCAT
CCTGCCTCAAAATCAAACACCAAATCAAAAAGCTCACATCTTGATCCAAA
AGAAGGTAGAGAGAATACACTGGGAAAGTCTTTGAAACCTCAAAGCTAAC
TCCAAGTGACAGTGACACCTCCTTAGCAGGGCCATAAATTCTAATCCTTC
CCCAAAGCCCACCAACTGGAGACCAAGTATTCAAAGATAAGAATCTATGC
AGTCCATTCTCCTTCAAACTACCACAGTAGGTTTTCTTAAAAAAAGAAAA
AAGAATATTTTAATTGATTGTGATTATTCAGTATTATTCATGAATAATCA
TGAACTACATGGCAGGACTATAAACTATTATTTTTTTTAAAGATTTATTT
ATTTATTTTATGTATGTGAGTACACTGTAGCTGTCTTCAGACACACCAGA
AGAGAGCATCAAATCCCATTACAGATGGTTGTGAGCCACCAAGTGGTTGC
TGGGAATTGAACTCAGGACCTCTGGAAGAACAGTCAGTTCTCTTAACCAC
TAAGCCATCTCTCCAGCCCCTATAAACTATTATTATATTTATAAAATATA
AATCCGTGAGTCTGTGCACCCCTGTGTGCACATGGATGGGACATCTTTGA
ACTGGATTATATCATACTTAGAAGAATACAAGATACTCTGTTTTGTCATT
TGGGTGAAAATATGGTCTGTTTATTTTGCAGGTATGACCTGACTTCTAGG
GAATGGCTTCCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGG
TCATTCTTTGGCATTACATAAGGTAAACTATCTCAACTCTTCACCAAGCA
AGAAGTTCAACTCTTCCTGTTGCTTTATGTCATTGAATACTATCGAGCTT
TGGTTTTAGTTGGTATAAGCTTTGTTTTGATGTCATGGAGGTATATAATT
CACCAAGTTGTCACCAAGTTGTAATTGGAAATTGAAGTTAGAACGATTTT
AATCCATGGTGTCTTGCATTTGGATACTCTGATCACAGTTAACAATGAAG
ATTAAATAGTGTCAGCAAGCCTATGCCCATTATCAAGTCTAGCATACTGC
ATGCGTGTGACTGAGTAGCCATTGTTATCTCCTTGTTTTGAGCGTATATT
GTAGAATGAGGCAACTGTATTTTCCACACCATTTTCGTTCTGTAACACGT
TTCATGTAGAGAAGGTGATTTAGAGAGGGGAAGAATGTGATTGTATTGGT
TGGTTCTTTCTCTATGCTATTCCTAGCAAGTCACCGAAGAGCTCATGTTA
CTCACACTTCTTAAGCTGGGATCACAATGAGATTGTGAACCACTCATTGT
TGTTTTCCAATATAATTTTTAAAAAGATGTATTTATTTTTATTTTATGTG
TGTGGGTGTTTTGCCTGCATGTATGCCTGTGTATACTGTTCCTCCAGAGG
TCAGAAGAGGATGGCATCAGAACTGGTGGCTGTTAGCTGCCATGTGGGTA
CTAGGAACTAAACCCGGGTCCTCTGCAAGAGCAGCAAGTGTTCATAAACT
CTCTCTCCAGCCCTAGAGTTGATTTCTTAATGGTTTTAAAAATCCTGTTT
ACATCTTTCTTATAGGATAAAATCTACATGTATGGAGGAAAAATTGATTC
AACAGGGAACGTGACCAATGAGCTGAGAGTATTTCATATTCATAATGAAT
CATGGGTATTGTTAACTCCGAAAGCTAAGGATCAGTATGCAGTGGTTGGA
CACTCAGCACACATTGTTACACTGGCATCTGGCCGTGTGGTCATGTTGGT
CATCTTCGGTCATTGCCCACTCTATGGATATATAAGCGTTGTGCAGGAAT
ATGACTTGGGTATGTATTTTTTCCAGTGGAGGCATCTTGAATATCATACT
GAGAACCCCTGCCCTTATTATTAGGACACCGTAACAAAATTCAGCATGAT

FIG. 3D(3)

CTTGATCCAGTACCTTGTCTTGAAATAGTATCAGTAGATAACTGGTGAGA
TTGAGGTTGTTGAAGTCCCTGTGCAACAGCTGTTTCTTACTTGTCAAGGT
CTAGTCTTGGCTTGGGAGGGGTTCTGAGGAAAGGGGTGTCAAAAAACCCA
AAAAGTCCAATTGTAGGTCCAAGCTGGCAGCTGTATATTGCATTAAGGAA
AGCTGAGGGAAATTTGGGATATTTATTTCATCTATTAGTCTACATCAAGC
AAGTCAAGCGCTCACAGTCAACGTTTGCACCCTCAAATTAGTAACAAAAG
AGGGGGAACTGAGGAGTCCAGCATGGTCCTGGTTGGGACAGAATGACATG
GTTCCAGCCCTGAGACAGGGGCAGCAGGTCCGGGCCTCCATGGATGTCAC
ACTATGGACATAAACCTGTTTGTATAATAATGTACATATTTCATGCTCCT
CTTCTGAGTAATGTCCTTCTGTTAATGTGAATGACTTCATGATAATCAGA
GCCAGTGTGAGTCTGGGAAGTAAATGGTGGGACCTTCAGGACAGCTCTTA
AGGCTGTGGAAAAGAACATGAGTTCAAAACCATATACTTCCTCAACTATA
CAAAAATAGAAGGATGCAATATGAATTGTATGAGGGGCTTCACAGATCTA
AAGGAACAAAAGCAGCTTCGCTGTGAGCCAACTTGTCAGAAAGATATTGA
GTAAGCAGTTAAAGAGATTTAGGGAGTGCTGATTGCTAGAGGAGGCCACC
CAGCTAAGTTTGTGCTTACAAAGGCAGACAAAGTCCTGAGTTCAGGGTGG
GCCTGGAACAGAGCAAGGTTAGTTAGACCTTGGTGTGGTAGAAATGGTAA
TTTCCAGACAGGATACCCAACTAGTTTTTGTGCTTAACAGAGGCAGGTAG
ATCTCTGAATTCTTTTGTAATGTTAAAAGGAAATGTGTGCTTGTTGTCTC
CCAAGGGGCCTGAGTCCCAGGATGCTGATTTATAGGAAACCTGGAGTAAC
TGGGTTTATGACCTGCAGGAGACGAGCTATCCAGAATGTTTTTTGCAATA
GCAAGAGAGAACTGCCTGGAGAACTGCCTTCAGCAAAGAATAGCAAGAGA
AAGCTGTCTAGAGAGAGAGCTGTCTGTAGAGAAAGCCGGTCAGAGAGAAA
GTAGACTGGAAAACTGTCTCCAGCTTGGACCCACAATTTGACTTTTTGTT
TTTGTTGACAAGTTGCCCTCCCCCAGAAACACCTTCCTCAGGACCCCTCC
CAAGCCAAGGCAGGGCCTTGGCCCTTCTTGTCAGCTTGCAAGGAGCCAAA
GATAGCATTAAATGCTTTGGATATCAAAATAAGCAAAATGCAAAACAGTA
AACACTCTAAAATAATTCTGGCTAGTCCTTTAAATATTAGGCCAGTGCAC
TGTTATTTTACCTTAATGTATAATCTTGTGTTACATTTTATTGTTTTTAT
TGTATAATAGGAATGTCAGAATTATAATTTTGTAACATTTGTTTGACATT
CCTGTGAAAATGCATCTAAAGATCATTAAAGTGCATCTGAAGATCATAAG
GACTCACTGAGGAGCACAGGGAATTAAGTGTCTGCTTAAGAGAACTTTGA
ATCTTTAATCTTTAGAATTTGTTTTAAAAATTTGAATCTTGCCAGTGTGG
TGGCGCATCCCTTTGGTCCCAGCACTCAAGGGGCAGAGGCAGGTGTATCT
CCATTAGTGTGAGGCCAGCCTGGTCTACAGAGCAAGTTCCAGGCCAGGCA
GGGTACACAGAGAAACCCTAGCTTAACAAAACAAAACAAAATATGAATCT
TTAAAAACTTGTTCTGTGAAAATTTCATACATGTATACAATATAGCTTGT
TCATATCCACCGCCATTCCTTCCAGCTCCTCTAGGCTTTCCCAGTGCATC
TCCTTCCTAGCCTTATGGCCTCCCTTTCAGGGTGAAGGTTAGCACACTGA
GTCCAGTTAGTGCTGATCCGATGCAGTCTTGTCTAGATGGTCTTCTTTAT
AATAAGGTGAAAGTATATCCTAAACTTCCGTCTTTTGCTCTAAGGTGTTT
AGACTTTAAACTAATGTTTAAATCGTTTAAATAATTTATTATTTCATAAG
AAGAGGAGCCTGCAACATTGACTTTAACTATTGTCTCTTATCCAGAAAAG
AACACATGGAGTATATTACATACTCAGGGTGCTCTTGTGCAAGGGGGTTA
TGGCCACAGTAGTGTTTATGATGACAGGACCAAGGCTCTGTACGTTCATG
GTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGATGACCTC
TACAGATACGATGTGGATACTCAGATGTGGTGGGTGTTTTCCTAGAGCTT
TCCCTTGGTAGTCTAGAATCTGCAGAGGCAATTGATTAAAAATACTGTGC
TATGGTTTGACTTTTGTTCAGCATTGTATGTAACAAAGTTAGGAGATCAA
TACAGTAATAGAGTTAAGGTACTAATGGTGCTGTTGCTGTCTGTTAGTGC

FIG. 3D(4)

TTAGTGCTTTAGACCTGATTCACTGAACTCTAGCAAGGTTTCCTCTCTTC
AGAATTCTCAGCAATAAAAGCTGTGCTGATTTTATCCATACTTAAAAAGC
ATATCCTTCCTTTTCTCTTTTTGGTGTTGGGGATCAAACCTTGTACATGA
ATAGGCTATACCATCTTTATCCATTTACATCACCAAACAGGATGCTCTCG
TGCCTATTTGATAGGGTTTTCACTCACTTCGAACTGAAACTTGGGTTGTA
AGAGTATGGTACTTTTAGCAAATGGAAATAAATTTGAGTTATGATGCAAT
TATAAAGCACTGGTCTCTCTGTATTTCCCTCCTCCTTCTACTCCCTCCCT
CTTCCTTTCTGACCCCCTCTCTCAACATACATTAGAGACCATGCTTTGAC
TGTCAATTTATGCTGTGCTGAAGATCAGGTCTTTAGTGGCTGTGAACCAC
GGAGCCTATGCAGTGGAAGTTCTGGTCTCTGGCTTTTGCCTTACTAATAA
AACACTGAGCATAAATTTTGATTTGTATTTCACAATTCTTACCTGGAATT
CTTAAGTGGAATTATGGAGCCATAGAGAATGAACATTTTAGGGCTTTTAA
TATAGTTTCCCGAAATTTTAACAGATTTTCATGATTGTTAAAGGAAGTGG
CTTACGTATAGGGGGAAATCAAGTATTGCACATTTGAATCTAAAGTTATA
AAGTAATTACATTTAAATTGGCAAATAAGTATTCTTTTAAAACTAACCTT
ATATTTATTATTTCTAAATAAACTCAAAAGGACCATTCTTAAGGACAGCC
GATTTTTCCGTTACTTGCATACAGCTGTGATAGTGAGTGGAACCATGCTG
GTGTTTGGAGGGAACACACACAATGACACTTCCATGAGCCACGGTGCCAA
ATGCTTCTCCTCAGACTTCATGGCTTATGACATTGGTAAGCTTTCCAAAG
ATGTTTTAGCTTCAGGAATATTTTCTTTGCTGATGGAAAGATCACTATGT
TAAAATAATTGCACCATTTAAAAGAAGTCCAGGTGGTAGAATTTGCATTT
AATTTGAGTAGGGTTACACATCTATTGAAAAGCATTATTTTGGATTAAAC
TACATTAAATTCTTTGTGAAATCACTCTTCTTAATTGCTTTAATTCTTTT
TTTAGGTTGAGTTAATTGGTATCTTCTTTCTTATAAGTGCCTTACATAGT
AGTGGTGGTAGTTGTAACCACCAGTGTTATGTTAAGTTTGATGGGATATG
CTGTTTCCTAGAAACCTGGTTTTACACATGCTGTTGATGTCAATATACAT
GTGGCCAGAAGAGGGCAGTGTCTGTTTATTCCTGGAAAATAAACATCAGC
TGCTCTGTTGTGTAAATATCACCCATGTGATGTTCTTTCTGTTTATTTGT
CTTTGCATTTTGAGACAGCCTCACTATGTAGTCTAATTGGCTGAAGCTCA
GTATATAGATCAAGGTGACCTTGAACTTAGAGAAATCCTCCTGCCTCTTC
TGAGTGCTAAGATTAAAGATGTGTACTACGAATGAAAAAAAAAAATGTGT
ACTACCACACCTGACTAGAGATTCATTTTAAAAATTATTCTTATTGTGAT
AAAATGCTCAGAATAACACTCACCATCTTAATGTTTTAAGTAGTTTAGAT
TTAAATATATTCCTAGTGTTATTCATGTTATAATACCATCTGCTTGCCGA
CTTCTTGTAAAACTGAAACTCTGCCCTTAAACAATAGTTCCTCTCTTCAT
CCCTCACTCCAGCCTCTTGAAATCATTTTCTATATCTCTATGATTTTGAC
TAGTCTAAATTAGGCATTTTTTAAAAAAAATATTTTGTTTACTTGTATGT
GTATGAGTGTTTTGCATGCATGTATGTTAAGCACACCATGTATATTCAGT
GCCCATAGAAGCCAAAAGTAGGCATAGATTCCCCAGAGCTGGAATTACAG
ACTTTTGTGAGCCACCATGTGGGTGCTGGATACTGTGCCCAAATCCTTTG
GAGGAATAGTGAGTCTTCTTAGCTGTTGAGCCATCTTGTCAGCCCTAGAT
GTTTGTTTTTAACAAACGTGTTTTTGCCAGCCATTGAGTTTTTAAATTGA
GAATGGGGGGTACACTATAGTTAGTCCTTAGCTTCAAGCTTGTGGAAGCA
GAAATGAGAAGACAATATAATCTTAACTCAGGAGGATTCTTGCTGGCTGA
AACAAAGATGTGAAATTACCTCCGAGCACTCCTAAGCCACTGGGGTGAGC
AGGGTGGTCTGGAGAGGCCTTGAAGAGAAGCTGTCTGAGCTTGTTCCTGG
GGACACTGGGAGTCAAATAGACCTCCTGGGCAGGGGGATTTAGTGCAGAC
AAGAGGCAGGAAAGTACATGTCAAATATTTAGGACTTTTGAACCGCTACC
TTTCTTTTGTCATGGTAACACAGAAGGTAGCAGGTGACTGTTAGACTAGA
ATGTTCAGATCTGATTCAGAGTGCCAGGGATCGTTGGTTGGTCTTGTGTA

Fig. 3D(5)

AAGTCTCACAAGTGATAGAATCATATGTGTGTCTTAGACTTTTTTTGTTG
TAGGTATTTTAGATTTTTCTTGTTTTTCCTTTTTGTAAGTCTGGCCCTCA
CACTATGGTCCAGGCAGGCTTAAGACTTATGGTAACCATCCTACTCTGCC
TTTATGGGCCACCATGACCAATTTAAGAAGCTCTCTTGGGTGGCATTGTG
ATAAGTGATCTGGAAGGGGCATATTGACAGTTAGCAGGCTGCTACTGCAG
AAGTCCTAATTAGGTTTGTATCAAGGCCATGGAAGGAGCAGTGACTTCTA
GTACCTGGCTGTTGTGTGTCTTGACAAAAATATAACTGCCCTTTCTTCCC
AAGTGTCTACTATGGACCACCTTTGCCAAAACTAAAAGCAGATTCAGAGA
AAAACATATCATGATTGCACATGGCTATAATCCCTGAACTTAGGAGGATG
AGAAATATGGCAAGATTGAGACCAGTCTGAACTATCTAGTAAGACCGTGT
CTTTAATAAAAATAGTAAAAATTATAAAATCAGGGAGTAGGATCTGGGAA
GAAGAGAATGAAGTAAGTGTGGGGCATATCCAATTGGAGATGTCTTTAGG
ACAGAGCTGATTGCTGAGAGGTGGTTGTAGGAGAGGTGAGTTATTGTGGG
GCATAAAAGATGAGCAAGAGTCAGAGACAGTTGGAGAACAGAGTCTGAAC
AAGAGTAGAGACTAAAGAGAGTGTCAGAGAAGCAGGGAGAAAATAGGTGA
GATTGATGACCTGTGAGATATGTTAATGGCCAGAAGAGTGGCTAAAAATG
ACTGGAGAATCCTTCAGACTTGTCAACAAAGAAATCCTTTAGCCTAATTT
AGGGTGCAGGCGGCTGAGGAAGGACATAGGTGAAATATGTGCTCTGTGTG
TTCATTTTTATTAAAGCTTATCTGCAAAGGCCTCAGATTTGCTGTGTACT
TGTAGCTGAGGCTCTTTTGAACTCCTGGTTCTCCTGGCTCCACCTTCCCA
AGTGCTAGGATTACAGATGTGTGCCCTAGTTAAAATAGCTGTATACCTAG
CATTAAAAATTTTAAGTTAGAAAATACTGTGGTGCTCCGGGGATGCATCT
CAGCAGTAGAGTGCTTGCCTGCTATACACAAGGCCCTGGGACTGATCCCT
AGCACCACAAATACTAAAGCAGACATTCTGGTAGGGAAAACTGGTAGACA
GCAGAGTGGTGACCATCAGGAGGGGGGTTGTGGGTGATGAATGACTACAT
TAATTAGAAGTTCTGTGCAGTATATTTATTTCATGCCCTGAAACATTGCT
GCTGCTGTTGCTGCTTTCCTTTACACATAATAACATAACTAAAAGACAGA
CAAGCATGTGGTATGAGGCTGTGGATGAGGCATTCTTTGTTTTCCTTTTT
TTTTTTTTTTTGAGACAGGGTTTCTCTGACCTGGCTGTCTGAAACTCAGT
AGGTAGAACAGGCTGGCCTTGAATGCACAGAGACCCTCCTGCTTCTGCCT
TCTGAGTGCTGGGTTCAAAATTTATGTTTTTTTTCTATAAAGACTGAGAGT
TCACATGGACTATATATGACAACCTACTCTGAAATGTGTTTTTCTCCCCC
TTAGCTTGTGACCGATGGTCAGTGCTTCCCAGACCTGAGCTCCATCATGA
TGTCAACAGATTTGGCCATTCAGCAGTCTTGTACAACAGGTAATTGGAAA
GCAAAGGCTCTATTACTGTCTTACATCTTATATTCATTTTTAATATCAAC
TTCCTAACAGTTGTATCTGAATGGTAAGAGGTTTGGGGAGAAAAAAGGAG
AGAAGGCAGTTCTAAGTGCACGATAAGGTAAGGGGAATAGGACTGGGAGG
TTATGGGGTCAAAGAGCAAGTCTGAAGTCTGCACTATATCCAGGTGTGTG
CTCAGGAATACTTTTCTGACCAGCAGAGCTCTTTTTCCATTTGCTCCAGG
AACCTTAGTCCTGTAAAGGACATGCAAAGGACTAGGGTTGTGGGCCAGCA
ATAGAGTGTTTATCTAGCTTGCACAAGATCCTGAGTTCTGACCTCAGCAT
TTTGCCTTCTGCAAACACAGCATTTGCCATAAGGGACATGCAGAATGGCC
ATTTTACCTAGTCACTTGAAAGTGTGCTTTAAGATTGAGAAACTTAACAG
CCTGCTGATGCTGACTTTTCTTATTTTGCTTCTGTTACTGCTTTCTGCTT
CTTTCTTTAATACTCTAATGCTTACATTATATAGTCCTACAGGTATTCAA
ATTTTCTGTTGGAGTTTCCTAATACAAGTAATTTAACTTGCATTAGGAAA
AGGATAAAAGTGCCATTCTGGAGTTGTGAAGAATGACCGTTTAGAAGCTA
GATAGTGGGGAAAGATGATATCTTTAATCATGTGATTATTTAGTGTTTTA
CAAGTATATAGGGGATTGTGGCAAGACCATTGTATGATTAGAGACTAAAG
TGGAAAGATTTTTTAAATATCTTGTTAACTTGAGTGTTATCTTAAATTAC

FIG. 3D (6)

AATCTGATGCTTTCCTTCAGAAAAAGCCCTAAATGCCTCTTGAGGTTTTC
ATCTGGCAAGTATCATGTCACCTGGCCTTGCTGGTGGAATCTGCCCCAGC
TCATGTGTGTTCTTAGTGTTCTCCTAGCACAGAGTTAGGCACGTGTGGGC
ATTTGCATACTAATGTATAGTAATAGTAACAATTGAATGAATTGTCTATT
AAAACATTCTTAAGTTTTACCCAAACACAGAGAGGTCGACAATTTGTCAT
AAAATGTAGTTTATCCATGAATCAAAATCAGGAATGACTGTCTGAACAGT
GTTTTTATTTTTTATTTTATTTTATTTTGTGTAATTTCTGTGATGTGTTT
GAATATCTCAGTTTTAGGCAGGATTGGAAATGTTAGAGGTTGGTAAGAGG
TCATGGTTGCAGTTTGATCATGAGAGAAATCGATGGCTCTCCCTTCATTG
CAGTGTTGTCAGTCAGCAGTGTGGGATCACCTATGTCTAACAGTTGTTCT
AATTGAGAGAGGATTACAGGAGGGAAAGCAGTGAGATTGTGAGGTGCTAG
ATGAGGAGATGGCATTTACCTAGCAGCCTTCTCTCCCGCCCTCCCATCAT
GTGACCTGAGAGATTCACAATTTCTGAAGATATCAGCTGTGCTTAGTTTA
AGCAATAGTTTTATTAACTAAATCCAACTTGATTCATGTTATTCCCAGGG
AACCAGTGGTAGGATTAAAAATGAATCCTAGTGTTCTTTTTGGTTATTGG
AATGTCAAGTTTTCAGACACTGTAACGAATACAGAGCCATACAATCACTA
TATTTATTTGGTCCTTTGTTGACTTAGAAAAATTGAAGCCCAGTTTAGGT
GAGCTACCAAATTTCTCATTGTGGATTAGTATTAAACTTGCGTGGAGTTG
TGGGATCTTGGAAGTGGGGGCTAAGCATCCGTGTTTGTCACAGCCCAGAA
GGAACAGATGAGGTTCCTTTTGAGGAGTCTTATGTCTTTATGAACTTGGA
CTTAGAAATATTTGATGTGTTTAATTCTGCTGTAGTTTTTTAAACTCTAG
CTAGTGAGCATCTTTTCACAGGAGCGCTTGAGTCTGACCTACAGCCATTG
TCTGTCTCTGGTGTGCATATTACAAATGCACTGGGAGCGTTTCTTGACCC
AAACATATAATTAGATTTTTCTTCTAAAAAGGTCTAGTTTGGGAAGGAAT
GAAAGGGATTAGAGAAATGTTGTGGGTTTGGTATTTATTTATTTATTTAT
TTATTTATTTAATGTATATGAATGATCTATCTTCATGTATACCTGCATGC
CAAAAGAGGACATCAGACTCATGATGGTGATGAACCATCATGTGGTTGCT
GGGAATTGAACTCAAGACCTCTGGAAAAACAGCTGGTGATCTTAACTGCT
GAGGCATCTCTCCAGCCCAATTGTTCTGTTTTAGTTTGAGGATGAACATC
TAATTTAGAGATGCCCTGCTTTTCCAAAAGTGAGTTTTAAACACTAATTT
CCATTGTCAGTGGATTGGTCTTTTAAGAATATAGGTAGTGGTGGCACACG
CCTTTAATCCCAGCACTTGGGAGGCAGAGGCAGGTGGATTTCTGAGTTCG
AGACCAGCCTGGTTTACAGAGTGAGTTCCAGGACAGCCAGGGATACACAG
AGAAACCCTGTCTCGAAAAGCAAACAAACAAAAACAAAACAAACAAACAA
AAACAAACAAAAAGAATATAGGTTGGAATAGGTTGGAAGCAGCCAATGAT
AGTGCATACCTTTAATCCCAGCACTTGAGAAGCAGAGGCAGGTGGAACTC
TGAGTTTGAGGCCAGCCTAGTTAGTCTACAGAGTATTTTCCTGGAGAGCC
AAGGCTATATATAGAAACCCTATCTTGAAAGGCCAAAAAAGGAGGAAAAA
AAAAAAAAAGAAAAGAAAAAAGAAAAAAAGAATGCAGGTTGGGCAGTCAG
GGTAAGTGTCTAAGGTAAGAGGAATTCTTCAAGGTGGAAAGTCATGAGTT
CTGCGCCAGCCTAGGCTACAGAGTACTGAAAGGGGAAGAGACTGTCCATG
TGTCAGACCCTCATTTCTCCAAAAGTCACATGACTATATTTTTTCTGTAT
TGCCCACTCTTCCATACATGCACCTAACAATAAATATTGAAGTTCACTCT
GTGGCACTATATCTATGTGATAGACTTCTAGAAAAGTGATTTAAAGTTCA
AAAGGTAAATACGTAGTTTTGTTTCAAGTTGCCAAAATCCCTTTAGTAGA
CTCCTACAATCTTACATGCCCAGTAGCAGTATAGAAGCTTGCTTGTTGCC
TTGAAGCCTCACCAATTCAAATATTAGGTAACATTTGTTACATTTTTCTT
TGTCAGCTGGATAGGTAATGAATGACACAACAATGTGTTCCCATTTTCTC
TGCATTACTAATTGAAGTCCTATCACCCACAGCAGACTGAAGAGTTCCTT
TAATATTTTATGGACTTTGACAAACCTAGGATTCATAGCTTCCATACAGA

FIG. 3D(7)

GAGGAATTTCACAAATAGCAAAGTTGGGCTGTTAGAAGAATAAAAAGAGA
ATTCTGAGTACAGCTTCTCAAAGAAGAGTCCCACGTAGGTGTCCTCTGGG
ATGTGCCTAGATGCAGGGTTATTGTACAGGAGCTCTTCTGTCTGCTCTCT
GATACTTGAGATTATAGGGTTGCAGGGAAATGCATTAGATGGCATTACAA
ACTGATAAGATAAAGTTAGGAGCTATCAGAGATTTAGGACATGGTTTTTC
TCTGTAAATGGGGCTTCTGGTGAGATTCCTAGAAAATGCTGTTTATAGCT
AGGAATGGGGTTATAGCTAGGAATGGGGAAAGACCTTAAGCAGTTGTGAG
CTGTGGTGGAATGCATGTGTTTTCAGTTTGCTAAGGCTTCCGGGAATACT
TTTCCTGTCGATAATTTTCTTTCACTCTCTTTGTAGCCTTCTTTGTATTA
AAATCCTCTCTGCTTGCTTTTGTGTGTGAATGTGTGTATGTGTGTGTTTG
TGTATGTGTGTGTATGCATGTGCATGTAGGTCCCTACATAGGACAGAACA
TATTTCCTGGAGTTATAGGTGCTTGTGAGCAGCCTTTTAGGGAACCAAAC
TCTGTCCTCTGGAAGAGTAGCCCCTTTAACTGCTGAGTCATTTCAGCCTC
AAGAATCTTCTCTTTTCCCTATTAGTAGAAGATGTCATCTTAGCTCTAGG
AACTACACCACCTCTGGCCTCAGTGGACACCCATTTACATATGCACATAC
AGCAGACAGACATATAACTAAAGATAAAATAAATCTTTTTAAAATGTCAT
TTCCCTGTGTACTAATTTTCCATGTACACACTCACAGGTAGATTTTTAAA
CTATTCTGAGTGATCACAAAGCAGAGCAGAAGGTGAAATTTGAGAGAATA
GATGATATTAGTGGATTTTGAGACCTTGAAAATAATGTCTCAGAGCATTA
AATTAATCACTCATGTATGTATGTATGTATATAAGTATGTATGCATGTAT
TATGTGGATGGGGGTGCTGTAGCACATGTGTGGAAGTCAGAGGACAACTT
TGTGAAGTCATGTTTCTCCTTCCATCTTTATATGGTTCCAGTGATTGAGC
TCAGATTGTCTACCTGTGTAGCAAGTGCCTTACCTGCTGACCTGTCGCAC
TAGCCCTCTCAGAGGACTTTTAATATTTGGAATATTTCTAACGATTGACA
GTCAAAAGTTTATTGTGAGCCAGGCACTTAAAATCCTAGCACTTGTGAGA
CACAAGATGGAGGTCAGTCCAGTCTACTGAGTTCTAGACCAGCAAGGGCT
ACACAGTGAAACCTGTCTCAAAAATTTCAAAAGCGGAGCTAGAGAAATTA
CCCAAGGAGCTAAAGGGAACTGCAACCCTATAGGTGGAACAACAATATGA
ACTAACCAGTACCTGGGAGCTCTTGTCTTTAGCTGCATATGTATCAAAAG
ATGGCCTAGTCGGCCATCACTGCAAAGAGAGGCCCATTGGACTTGCAAAC
TTTATATGCCCCAGTACAGGGGAACGCCAGGGCCAAAAAGGGGGAGTGGG
TGGGTAGGGGATTGGGGGGGTGGGTATGGGGAACCTTTGGGATAGCATTG
AAAATGTAAACGAGGAAAATACCTAATAATAAAAAAAAGAAATGATATCA
GAAAAAAATAAAAAAATAAAAAATAAAATAAAATAAAATTTCAAAAGCAA
CAACTCAAACCAGCCCTACGTCGTGCCTCTGAGTTCTCAGTAAATTCCTT
CTCTCTCTCCTCTCAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTC
CTCCTCAGTGACGTCTTGGTCTTTACCTCGGAGCAGTGCGATGCACACCG
CAGTGAAGCTGCTTGTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGG
ACACACAGTCGTCTCGATGTACCTCCTGGGAGTTGGCAACTGAAGAACAA
GCAGAAAAGTTAAAATCAGAGTGTTTTTCTAAAAGAAGTATGTTTTTTCT
CTACTTAGAATTTAAAAATCTAATTTTATCTGAATTGTGAAGGAACCTAG
TCTCTGTACTTTCCTGTTCACCTTACTCTCTAGTTATTTCTTAATAAAAA
AATACACAAGATCTTTGGATGGGAGGAAGCATGTGGCTCCTGGAAGCTGT
TAGCAGGTAATAAGTTGTCTTTGAATTACACAGGCTTTGTGTACCAACTC
CTGGTCTGGCTGCAGGTGATCTGAAGCCATAGCACAATGAAATTTGTTTT
CATTTTGGTTTTATGAGACAGGGTCTTGCTCTATAGCTCATACTGGTCAA
GCTCCTTGTCAGCCTCCTCCTTCAGCCTCTTGAATGCTGGGGTTATAGGC
ATGCATCACTGGCCCTACTTGGGAAATATTTTGATGACAGACATGCTATA
TATTTCTTTGTTCAGTTTAGTAGCCACTAGCAATCTGTTATTATTAGATA
TTTGAAATGTGGCTATGTAACTAAGGGGCTAACTGTTTTCTTTTCTTTAG

FIG. 3D (8)

TGTATGTAGTGAGGCAGATGTAGTAGCACACGCCTGCAATCCAGACACTC
ACGAGGCTGAGGCAAGAGGCAGTTCTAGGCCAGCCTGGGCTGTGTAATGA
GACCTTGTCTCAAGAGCCAAAACATCAACAATAAAAGAACAGTATGTGGC
TATTGGCTGTTATGTTGATGATGAAGGTCTAGTGTTAAGGATAAGAGCCT
CTAATGGTATGATCACATATAGCAAATTGCTCTGGTAGACAGCAGAGAGC
TGCTGTTCTTGAAAAGTATTTCCAGCCCCCTTTAGCTGTATATAGCAAGC
AGTACAGCATAACAGACAAACTATGGTCCCTTCTTCTAGAGCCCCTGGCG
TGCTCTTGTTATTTTTCTCTCCTTTGCTACTTGCTTAGTGGTTGCTCTGA
GCACCACTTCACCAACTCAGCGAAGTAACGTGCAAAAATGTTTGGAAAAT
AAGAATGCCTCCAAGATATTTGTCCATATCAATCTTTAAAGTATGAAACT
ACTTCCTTATCTAGTTGTTGCAGTTACATGAGAGTTATATTAGGCAGAGA
CTACTTCTGTTTTTCTGGTATGTGTTAAATAAAGTTGTGCAGGGACATAA
AGCTCCTGAGGCTGTGCTGTTGATTAGAATTTTGGTTCATTTATGGAAAA
CAGCTTACCAGAACCTGGTAGGATTCATAATTCTCCCGAAACAGTTAGAA
TTGGTAGAATAACCAAAATTTAAAGTTAAGCTTAAATATACAGTGCATTG
GAAATAATATTATCTTCTGAGGTTCAGTATGAGCCCATTAGTTTACCTCA
CTTTCTGGGTAGACCTAATCCTGTCAGAGTAAACTTGGCAAGAAAAGCAG
CCTACATGAAAACTGATCAGGCAGGGAAGTTTCTGTGGCCTCTCTTCCTG
CTTGTGTATGTCATATTCATGAAATGATTTATAGATGGCAACATGGCTTT
TAGCTTCTTGTTTGGGGATTTAATGAGAATTATGTTAGGTCTACAAAGAG
TGGAAGTTGTGAAATCCACAGGTTTGGAGTCACATGAGTATATAGAGTTC
GAGTTAGCAAGTGCCTCCTGTGGGGTTGTGGGTCACTGGGTATACCTGCA
CCCAGGTAGGCCTTGCATTTGTAACAAGGACAAATGTATTGGTCTCTCAT
ATTGCTTTCTTAGGCTTCTGCACAGCTTCTGGTGTTAATTCTGTTGCTAG
TTGATGTTTGTCGTGGGAAGAAAAGCATCCATTACTTCTTAGAAGCTATA
AAATTAACAGACCTTTGCTTTTCACTTTCTGGACACTATGGGAGGACAGT
TATAAAACAGTGTTTCTCGGATTGTCTGCTTATATCTGTTTTATTTTAAC
CTAAACATGGCACTGCTTTTTTCCTTTCAGTTTGACTATACACTTTGCTT
CCTGACTATTGTTAGGAGCTTTCCTACCTCAGATTATACATAAGAGAGGC
TGCCGCATAGTTGATGGGTTTGTCTTCTCTCTGTAGCCCTTGACCATGAC
AGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATACCAATGA
CTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAACCACAGCTGCACAG
AAGGCCAGGTCAGATGCTGTTTTTCACGGATTTTAGGGAATAGAAAAATG
CTAGATGAGTGTGAGTGTAGGGCAAATAATGAGTAGAGTTCTTTTTAAAA
TGGGATATCGATTTGAATTCTACTGTTGCTCAGGTTTTCTCTTAGGAAGG
GATGCTATATACATCCTGATTCCAAGGATCGCTCCTGCTGCTGAGGTCTT
TGTGCAGTGTTTCCGAAAGCATGTTTTACAGAATGCCCTTGGCCCATATC
TGACTCAGCATGACATCTGGGCTAATCATGTATGATTTGTTATAGGTGAT
AATAGGCTATGAGTAAGGTGATCCAGCTTTTGCTGTCTTTGATGGCTTAT
GACATTTTTTTCTCAAAGTTTAATGCATTTCATAAGAAATAAGACTTGAG
ATTGCTATGGTGGGCACGGGCTGGGAGGAGCTCTGGAAAAGCAGCAGGTT
CAGCTTTCACGTTTTACAGATAAGCATTGGCTGAGGCTTGGTGGTGCCAG
TGGTTCCGTTGGGCTGCTAGCTTGCCAGCTAAAAGCATGTTAGTGAGAAT
ACACACTGTGGTATTCACATTGCAGTGCTGCTTCCTGTTCATTCTAATTC
TATCATTCATCCATCTACCTATCTCTATCTATCTATCTATCTATCTATCT
ATCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCACT
TATCTAATTCTATCTGTCTGTCTGTCTGTCTTTCTGTCTATCTATCCTCC
ATCTAATTCTATCTATCTGTCCACCTATCTATCATCTAATTTTATACATC
CATCCATCTATCCATCTATCTGTCTGTCTATCATATATGTAATTCTAACC
ATTCATCTATCTATCCACTTATCTGTCTGTCATCTAATTCCATCCATTTA

FIG. 3D(9)

TGTATCTATCTATATATCTAATTCTATCTATTCAATTCTTTTTCTTTTTTTT
CTATCTTTCTTTCTGCAGTTACCATTCTCAGTTAATTCTCACTGAGTTAT
TTGTGTGAATAACAAAACACTTCTCCCCTGTGTTCCAGATCTCCATTGCC
AAGTATGAGAGTTGCCCCAAGGATAACCCCATGTACTACTGCAATAAGAA
AACCAGCTGCAGGAGCTGTGCCCTAGACCAGAACTGCCAGTGGGAGCCCC
GGAATCAAGAGTGCATCGCCCTGCCGGGTAGGCCTTGCACAGGGATGTCC
TCTATAAGGTCCAAGCTTGGTCCTCCCTCCTCAGATCAAGGTGGACCTAG
GAACAAGATTGCTTATTCTGTCTATTTAGCCCTCTCACTATTGGGGGGGG
GGGGGGGCGATATTTTGTATGTTTTTAACTTAAATGTGGTTTTTATGTAT
GTATTTACTAGCCTTTGAAAGAAAGTGAAGTGTCAGCTCATGTTCTGGAG
AATTGGGGGGTAGCTTAGATCCATGTTACAAACTGTGTCCCACTGTCCTT
CCTTCTGCTGTGAAGGAGAACCTGGCACTAGAGCTCTGTGGTCTCAGCAG
CAGTCAGGAACCTGCAGGAAGCACTTACTGACAGTTGTGTGAGAAGAGAT
TTCTGTACCAGCATCATCTCCCATGTGACCTTCCTTCCCGACTATTTCAG
CAGAGGTTGTTCAGGGTATTAACTTAGGTCCTGAGGCCAGCTAGCCCTGA
CTAAATCTCTATGATGTATTTGCTTGATCAGGATATCCAGGAAGGGGAGC
TTCTGTGCTCTCCAACATCGAGGTTTGAGGGGAAGTTGGTCTGACTCTTT
TGAAAGCATTTTATTTAGTTTGCTGAATGGGCTTTAGTTTAGCCAGTGTT
CTATTGCTGTGAAGAGATACCATTTCCAACGTGTAACTTTTATGAAAGGA
AACATTTAAGTGGGGGCTTGCAGTCTCAGAAGCTATTATCATCATGACAG
GGAGCATAGAGGCACAAAGGCAGGCATTAGAGTGGTAGCTGAGAGCTACA
TCCTCATCTGTGAGCAGAGGCAGACAAGGTGTGAAAAAGACAGAACCTGG
CCTGGGCTTTTGAGACCTCAAAGTCTACCACCCCCAGTGAGACACTTCCT
CCAACAGCTCCTGCAACAAAGCTCCATCCCCCGATCCTTCTCCAGTCCTG
CCACTCCCTGGTGAATGAGCACTCACATATATGAGCCTATGGGGGTCATT
CTTACTCAAGCCACTACAGGCTTTGTTTTGTGTCTCAGACTTTATGTCAA
TAGAATACCTAGACACCTTGTTACAAGACAGGCCTGGAAAGCCTGCAGTG
CTGACTCCCTGCCAGTAGCACATTCTGAGGAGCAAGTCCCTTAAGTCGCT
TACCTGCTCTTACATTACGCCTTTCCCTGACCATTTAGTGAGCACTGTTG
GTGTCCCCAACCTGAACCTGGTTCTGGGGAAACACTTGCTTATTCACTTC
CGTGCTAATGGCCAGGGAGCAAGCATGCTTTCATGCAACACTGTGAGTTC
AGTACAACCACAGGAGGAGATTGCAGACTTCCTTCGTGTACTGTATCACT
ATGAGGTTTTCCAAACCAGTCTCCCTTTCACCTCATTTTTTGGCATGCCT
TATGTACTTGCTTATACTTTCTATCTTATGACATGAAAACAGAGTGGCAT
TTGGAGGCTTAAATTTATCACATTCCCAATTCAATTCCATTTTCAGTTTA
CTCTTTCTGTATATACATCAGTGTGCAGATAAATATCTCTTTGTGTGAGC
ATTGGAGGCCAGAGGTTAACCTCTGGTATATTCTTCCTCTATCACTCTTC
ACAGGGTCCTTTGATGAATGTGGAGCTCACTGATTACATAGACTAGCTGA
CTCAACCCTCAGGCCTCATAACCCTGCCTCTAGCCCTCAGATGAGATTAC
AAGCAAGCAAAACTACGCCTGGCCTTTTATGTGGGTGCTTGGAATTTGAA
CTGGGTACTTATGCTTGACACAAGTATTTTATCCACTGAACCATCTCCCA
AGCCTCCATTTGCAGTTTTTTACCTCACCCTTCCAATATATATATTTATT
TGTATGCCCTTTGTTCAAGATTTTAGTCACCTTTTACATTTTTCTTCAAA
AATAATTGCACCAATTTCTTAATAATGGCACCCAAAAGTAGGAACATTAG
CCTAGAGTATACCCTGTGAGCCAGGAAATGTGACTGGTGAGACTTGTAAA
AGGGTCTTTTTATTCTGGCCCTCAGCGGAGGCTCAGCAGTGGAGCATGCA
TGCTGTTCCTCTGGAGGACCCGAGGTCCCCAGGGGCCAGGTCACAACCAC
TTGTAACTTTAACTCTGATCTAATGCCCTCTATGGCTTTTGTGCTATAGT
CTCTTGCACTAACCCACACTCAAGGCACACATACACACATTCTTTAAAAG
ATAAATTATTTTATTTTCAAAGGTTTTTTTCTGCATATAGAAGTTAATAA

FIG. 3D (10)

TTTGTCTGTTATGCTCACCAGATCCTAACAAAGCACCTGAAATTCAAATC
AGGATGAGTTCAGATGTTCAGTATTTTGAACTAGTAAACCGAACTGCATA
ATTCCTAAAACTTTGTTTTCTTTCCTCTTCCCCTTTAAAAAAGAAAATAT
CTGTGGCAATGGCTGGCATTTGGTTGGAAACTCGTGTCTGAAAATCACTA
CTGCTAAGGAGAATTATGACAATGCTAAATTGTCCTGTAGGAACCACAAT
GCCTTTTTGGCTTCCCTCACATCCCAGAAGAAGGTGGAGTTTGTCCTTAA
GCAGCTTCGATTAATGCAATCATCTCAAAGTATGGTGAGTTAATGTGTTC
AGAACTTTGGTTTCTAGGGCACAACAGCAGCTCTTATGTAGAAGGCCACA
GTTGTATGTTATTTGCCTGGTAAGAGAAAGAATTACAATAAATGATTAAT
AATATACTGTGGGCCTCTATTTCAGAGGCTCTTCTTTTGATACCTTTCTT
CTTGTCTTAAAAAGTTCAGTACTTTGCATATTTTATTAGTTGTTATTATT
AAGTAAATTATAAGGTATGAACATATGGAATGAATGGTAATATGTGTACA
TATTCTGGTGACATCAGATTATTTTGTACTTGATTTATATCTAGATTCTG
CTTGGGAAAAGGGAGAGTAAAATGTTAGTTACCTAGGTGTCATTAAAGCC
ATCTACAGCCCCTGGAGGTATTATTATAGCACATAGTGTAATCGTCAGTA
AGAAATGTAAAATCTGCCCAGGTTTTATAGCCTTCTTCCTAAGGCTTCTG
AACTCAGAAAGTTCTCTTACTCTAGAGCCAAACTCTCAAATGGCTTGTAG
TTACTATATAGTCTCATTTGGTATTTTTCTTGGTAAGTCTAATTCTAAGA
CTTGTGATTTGACTGTGATGCTTCAGTCAATTAGATATTCACAGAGCAGC
TTTTCTGTCTATGCTGGCTGTGGTACAGAGAGATGTGAGGGACATGTTTT
TGTCTAGCCAGGAGAAGACAGAATGCAGCTCAGCATCTCTCATTTGGCAC
CACCTTCATGTGATGGGATGCCGGTATGGTGTGGGTCCTGGTTGTTAAAT
CTCAGGAAGTCCATATATCCAGAAATGACCTCAACTATAGGTGGATTTCT
GGCAATTAGGTAAAAGTCAGCATTCCTTGGGCACTTGGGAAACTGGTTAC
CATCTGCATAAAGGAGTCATTTCCCTTCTATCTGGCAGAAGGGACATATG
GCTATCTATTGTGCCTGTCAGCATGGAAGCACATGCTAGTCTCCAGGTCC
CCCCAATATCACAAGTACCTATAGCAGTGAATTAGTTAAACTGATTTGGC
TCCCAATGGGTCAAGTACAGCTGCACCTGCCCAAGAGCTCTTTGGGTTTG
CAAATGAGAGACACATAGTTAATTTTTATATGCTTTGACTAGTTCAGTTG
CTGGACATTTCTAATCCTCCCTGCAGTAGCATACATTAACCCCTCCAACT
TTCCTGAGTCAACTTACTAACTCAACATTTCATCTCTGACACCCCAGACC
TAATGGCAGAGTGGCCCTTAGAGCCACTTTCCCAATTTTTTTTTTATCAG
ATATTTTCTTTATTTTCATTTCCAATGTCCCCTTTCCTAGTTTCCCTGTC
CTCTCCCCCTGCTCCCCAACCCACCCACTCCCTCTTCCTGGCCTTGGCAT
TCCCCTATACTGGGGCATAGAGCCTTCACAGGACCAAGGACCTCTCCTCC
CATTGATGACCGACTAGGCCATCCTCTGCTGAATATACAGCTAGCACCAC
GAGTCCCACCATGTGTTTTCTTTGATTGGTGGTTTAGTCTCAGGGAGCTC
TGGGGTACTGGTTAGTTCATATTGGTGTTCACTTTCCCAAATTCTTACAT
GGCTGGTTTAGTTCTTTCCTGCAGCTCTTAGGTCTAATCCCTTTCCTTCC
TCTGTCATGGTGATTGCCTTCCTCTCCTATCTCAGTTCCTTGCCTGCTCA
ATCTAAAAGTCCCACCTCCATCTTTCTGCCCAGCCACTGGCTGTATGCAG
TTCTTTATTATCAGTTGAAGCCAGCTAGGGGCAGAGACCTTCAGGTCTGT
AAGTGCTTTGGGGAGCAGAATTAAGACAAAGCATTAGAACCAATTCCCAA
CAAGTACCTGCTATACATTTCAAAGTCCATATTAGTCTCCTGGGTCTTCC
CTTCCCCAGCTACTTGTCCTCCTTGTAATCCAAATGACAAGCTTTTTCAC
ACATCTCTTTATCTCACATTTCCCTAGCCCTGGCCATGTCCACTTGTTCT
TTTACTCTCTGCTCTGCTCTCTTTCCAATGCCTCTGGATATTTTCTCTCT
CTTATTCACAATAAAAACCAAACCAAACCAAACAAAAAACCTTACCCTAA
TAATGGAGTGGTCACGCCTGAGGTTTCCTTACTGCTCCCCCTTGCACACG
TCTTGTGTCTGACACACTGGCAGGCTTTTATTAGCAGCAGGCTCTAGGAG

FIG. 3D(11)

CTGAGAGAAGCAGCAGGCACCTCTGAGGTGGTAGTTACTAGAGTGATTAG
AACAGACAGTGGAGACGTGGCTGGAAATATGGACTCTGGTGTTTGGAGCC
AAGTATGGTAGGCGGCAGAAGCCAGCAGAAGCATGATCCACACCTTCACC
AGGTTGCTTCCATTGGGAAAGGCTGGACCCCTTGGGAAGGGGTCCCTTTG
TGCCTTCCTAGGTGTTCGGAGCCAGGTGTGTGAGGGATACAGTAAAGGGA
CTGACTGCATGACTGCTCCATTAGGGTGAAGGGTTTTGTTGTGAATAGGA
GAAACAAAATGTGCAGAGGCATCTGGGAGAGAGCAGAGCAGAGTGAAAAG
GAAGCAGTGTAGGCATGGTCAGGGCTAGGGACAGCGGAGACAGCAAGATA
GCGAGTGGGTGATAAGGTGAGAGAGAGTGTGTGTGTGCGGTGCACACATC
ACGTGCATTATAAGGAGGCTGAGTAGCTAGCTGGGGGGGAGGGAAGGGCCA
GAAAACTAGCATGCACTCTGAAACGGGTACTTGTGATGCTGAGGGAGCTT
GGGGGAGAAGGGCATGCCTCAAGACCAGAAGAGGGAGTTGGAGTTACAGT
TTGTAAGATGCCTAATTTGAATGCTGAGATCCAAACTCTGATCCTTTGGC
TGAACATCATATCTGCTGAGCCATCTCTCCAGCCCCTAGAAAGGTGGTGA
TGGTGGTTGTTCTTGTTTTGTTTTATTTTGTTTAAATGGGGAGCCAGGTA
CAGTACATCATGCCTTTAATCCCAGCAGGAGATTCAGGAGATAGAGACAG
GTAGATCTCTTTGAGTTCAAGGGCACCTTGGTGTGTATAGGAAATTCCAT
CCACCCAGGGCTACAGAAGGGTACCTTGTCTTTAAAAAAAAAAAAAAGAA
AGAAAGAAAGAAAAAGAAAAAAAAAAGAGAATGAAATTTCAGAGTTATGC
AAGATAGGAGCTCAGTGGTAGAGTGTGTGCCCAGGAAGTGCTGGGTTTGA
CTCCTCAGAACAACAGCAGGGGCAGAAACTAGTCTACAGGTTCATGAGTG
GTGTTTTGTTTTGTTTTACATAAAATGTGTTGAATTAGATAAGTAGATAA
AATGTGACTCATACACAGATAAATAGATAAAATGTGATACATGTACCTGT
ACATAGAAGATTATGATCTCACCTTTAAAAAGGAGGAAATAGAGAGTTTT
GGTAGTTACACCACAGGAAAACTGGAAAAGAAAATGTATATATGAGGCTG
TGCCCCATGGCTAAAGGAACATGTTTTTAAGTCATTTGAATTCACCAAAC
AGTTTTAGGTAATGATATATGGTTTTGCATACAACCAGTATTTTATAAAT
ATTAGCAAGGTCACATCATTTATGAACCAACATTTAAACTAAATTTGTAA
ATCATCATTTCTTTATAGCACTTGTCATAGAACATAAGTAGTTTAAAATG
TGATTATTGCTTTGCTCTTGATGTCTGAAAATCTTCATGTATTCTCTTCT
TTGAGCCATTTTTATGCTTTGCAGTACTGGATGCATATTGAAGTGATCAC
TTATTTTAATCTACCTTGCCTGAGTTTGGGGAATAGATGGTTTCCACATG
TCTGTGGGTTATGCCTAAGCTAGTGGTTTTTATGTTAGAGCTTGTTTTGG
GGAAGGCACTGGTTGCATTCATAGCTGTGTTTCTTTTGCCTGTAGTCCAA
GCTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTGTCTTACT
GGTGCTGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATG
CCATCTGAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGTCAGAGCCTAG
TACTCGGGGATTAAAGGCTGCAACCTGCATCAACCCTCTCAATG***GCAGC***G
TCTGTGAAAGGCCTGGTAAGGACATGGGTGCATATAGTGCTCCAGGAGGA
GCCAAGACAGCAAAGGAGGCACAGCTGAATGAGCGCTGAGGTGATGAAGT
ACTTATGGCAGCAGGGAGAGGAGCACCAATTTAGGCATATGTATTTCAAA
CAGAACCCGATTCCAGATAGTCTTTCTTGGCCTCTGACTGCTTTAAGCCA
TACTGAAAACCAAAAATAAAATTGCTGAAAGAACCCAGTTTATATTGAGC
TGCACTGTTTCGTTGGTCTCAAAGTGTTGAGAATTGTTCTAGAAGATTAT
TTCCTTGGTGTTGGCAGAGAAGTGCTATGGAGGAAACAACAACCTGAAAC
CAAAGAAACATTTAGAAAAGCAGCAAGTCAGGACACTATTCAGACACTGC
TGGGGTGGGGGGAGAGGGGCATGGCCAAAGAAGCCGACAGAGCCAACACC
AGGCTGTGGCAATGTCCTGCGCTGAGGTTAAGGTTAGACTCCATGAGGCC
AGGCCCAGAACAGCCATACACAAATGAGGACTCCAAAACAAGAGGTGCAA
GTGTAGTGGAGACTCCATCCCTGCAGGTCCTGTTTCAGGAAATGATTGTA

FIG. 3D(12)

CTTTGCCTGAGTAATACAGCCTAGGAGCTACTTTCTGATAGGGTTTTTTA
AATACTTACAAAGAATTATTTATCTTTAATCATGTGGTTTTGTATGTGTG
TGCTTGCACATGCAGTGCTTGTGAGAGAGAGTATGTGTGAGAGCATGCAT
GTATGAGAGTGTGAGAATATATGTGAGAGAGTGTGAGTGCATGTGTGCGT
GTGTGCATCTGTGTGTACAGGTGTGTGTACATGCATGTGTGTATAAGAGT
ATGTGAGAGTGTGGGTGTGTGTGTGAGAGTATGTGAGAATATATGTATGA
GTGTGTGTGAGTATGAGTGTATGTGCGTGCCTGCATGTGTGTGTGTGTGT
GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTAGAAGTGGCCTTGGAAAACA
GAGTTGTCAGATCTCTTAGAGATATAGTTGCAGTTGGTTGTGAGCCATCT
CATATGAGCGCTGGAAGTTGAAATTGGGTTCTCTGGAATCCTCTGGGTTC
CTTGTTGAAGCCTGAATATTTTGATAAATATTTATGTCATTATCCCTCAA
AATTGTAAATGTAGAATTTAACAAACTCAGGTCTTGAGTCATCTTTGTCC
CAAGGTTTGTTTGTTTGGTTTTTTGTTCCCCCACCTTTTCTTCAGTGCTT
TTAAAAAAGAGAGTCCATTTTTTCCTAAATGTTTAAATACAGTTGAGGAA
TAGAACATCTGACTCCAATTTCCTGGGTTTCCCTCCATGTAGTGTAGTGC
TGACCTGATTTCAGTGTGCATTGAAAACTTTGATCACTTGGAAGGCAGCT
ATGCTCACCACTATACTACCAATGTCTGCAATCCTATAGGAGAAACAACA
ATATGAACTAACTAGTACCCCCCAGAGCTGTGTCTCTAGTTGCATATGTA
GCAGAGGATGGCCTAGTCAGCCATCATTGGGAGGAGAGGCCCTTGGTATT
GCGAAGATCATATGCCCCAGTACAGGGGAATGCCAGGACCAGGAAGCAAG
AGTGGGTGGGTTGGGGAGCAGTGCGGGGGGGGGGGGTATAGGGGGTTTTG
GGGATAGCATTTGAAATGTAAATGAAGAAAATAACTAATAAAAATTGCCT
TAAAAAAAAACAAAAAAGAAAAGTTTTTGATCTTAGCTGACCAGTGTCTC
TTTGGGTCTTAATTTCCAGCAAACCACAGTGCCAAGCAGTGCCGGACACC
ATGTGCCCTGCGGACAGCGTGTGGCGAGTGCACTAGCAGCAGCTCGGAGT
GCATGTGGTGCAGTAACATGAAGCAGTGTGTGGACTCCAATGCCTACGTG
GCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGCTG
CCCACGTAAGTGGAAGGAGCTTTTGAACATTTGCAGGCAAGTTGGGCTTG
ACTTTCTGCTCAAGTCCATGCAGAAGCTGGTCGGGCCGGCCCTTCCAGAT
TAACATGTATGTATAGAATGCAGCACAGTGTTCCATGCAGTAAATCAGTT
ACATCAAGGAGAAGGCACAGGGTACAGAAATACCTTTTCTTCTTCAGGGT
AATATTATAATTCAATCTGTATAATGTTTCTACATCTTAATCTACCAGTA
TGTAAAGTGCTTTCTAGTAGAGGCCTCCCCAGCTCCCTTTTTCATCCAAC
ATCCTGATATTAAAAGGTTGGAAAAGTCCCTGTTATATATTATGTAAAAT
GTGGGGCCCTTTAAATTATTTCAGTTCAATAATCACTATAGGGTACTATT
TTTAATTCATGGAAGTTAAATCATCTGTTAAAAGAAAAGGTAATAACAGT
AAATTCAAATCTTGTGATAGTGAATTACAAGTTGGATTGTTTTGCCTTGT
TTTTTAATAGCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTT
GGAGCAGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAG
GAAAATGTATTGAGGGCAGCTATAAAGGACCTGTGAAGATGCCGTCACAG
GCCTCTGCAGGAAATGTGTATCCACAGCCCCTTCTGAACTCCAGCATGTG
TCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGGTAAGA
TGCCTGTGTATCCTAGTTCAAATCTCGTACATAAACTAGACGCCCAGATC
CCTTGGCTCACTTGTTTTCTTGACTGTGTTTGAGTTCTTTCTGTGTTCTG
CATCACCTTGTTGGATCATAGCTGGCAAAGGTGCTCTCCTTTCTGTGGGC
TTTTTCTTTACTTGATTGATTGTTTCTTTGGTTGCACAGAAGCTTTTTAG
CTTTCTGAAGTCCCATTTGCCAGTTGTCCTTAATTCCTGGGCGAGTAGAA
GCCTCATAAAAAAAAGTTCCTTCCTACACATGTATCATGTAGGGCACTGC
CTATGTTTTATTCCAGAAGTTTCAGAGGTTCGGGTTATGTCTTTGATCCA
TTTAGGGTTACTTTTTGTGAAAGGTAATGGACACAGTTCTGTTTCATTCA

FIG. 3D(13)

TTATTCTACATGTGGACATCTACTTTTCCCAGCACCAGTTTTGAAGATGT
TATCTTTTCTGCAGGTTGTTTGTTTGCTTGTTTTGTCTCTTCAGAAAATC
CCAGATGGCGGTAGCTGTGAGTGCTTAGGCTTGGCCTACCTGTTTCATTA
TGTTGGCTTGCATGTCTGTTTTGTGCAGTGCCACCATATTGTCTTAATTG
CTATAGCTCTGCAATCTATCTTGACATCTGTGTTGGCAATCCTGCAGTTT
CGACCCTTCTGCTCAGCAGTGCTTTGGCCATCTGGGGTCTTTTCTGGGTT
CATAATGAATTTTAGGATTTTTTTTTCTATTTCTGAGAAAGTATTGTTGA
TATTTTGATTGCGATTGAATTGAATCTGTAAATTGCTTTTGGTAGAATGG
TCATTTTCACAATATTAATTTTACTGATCCATGAACATAGGATGACTCCA
GTCTCTCATGTCTCCCTATAGCCCTGTCTTAAGAGATTTGGAGTCTTCAT
TGTAGAAGTCCTTCACCTCCTTGGTTAAGTTTATTTCTAGATATTGTATT
GTCTTTGGTATTATAAATGGTAGTATGTCCATGATCTTGTTCTCAGTGTT
TTTTTAGTTTAGTTTTTTTTTAATTTATGTGTATGAGTGTTTGTTTTATAT
ATGTGTATATGTGCATTCATGTCCTCTGGGCATCAGATCCCCTGGGACTG
GATTTACAGACAGCTTTGAGCTGCCTGTAGGTGCTGAGAATTGAACCCAG
GTCCTCTGCAAGAACAGCCAGTGCTCCTACTCCCCAGCCCCAGAAGTACT
AATTTTTAAGAGCTGATTTTCTACCTTTGCTGACATTGTTGATTGTTTCT
AGAAGTTTAGTGATAGAGTTTTTGAGATTTCTTATATATCTTATGTTATC
TGTAAAAAGGGATAATTTGACTCCTTTTCCTATTTATATCCTTTTATTTC
TTTCATTTGCCATATTGTTCTAGCTAGTGCTTCCCGCTCAGTATTGAAAA
GAGTGGTGATTGTGAACAGCTTTTCTTATTTCTTATTTTAATGGGATTAT
TCACCCATTTAAGATAATGTTGGTTATGGGTTTGTCATACACAGCCCTTC
TTATATTGAGGTATGTTCCTTCCAGTCCTGTTCTCTCTAGGACTTTTTTT
TTTTTAATCAAGAAAGCATATTGGGTTTTTTGTTGTTGTTATTTTGTTTT
GTTTTTCTAGACAAGGTTTCTCTGTGCAGCCCTGGCTGTCCTGGAATTCA
CTCTGTAGACCAGGCTGGCCTTGAACTCAGAAATCCACCTGCCTCTGCCT
CCCGAGTGCTGGGATTAAAGGCGTGCACCACCACTGCCTGGCACATGTTG
GTTATTTTGCAAGCCCTTTCTACATCTACTAAGATGAGCATGTGGTTTCA
TCTTTGTCTGTTTATATTGTCTGTTGTATTTATTGACTTATGTGTGTTGA
GCCAACCTGAAGTTCTGGGATAAAACCCACATGCTTTGGATGATTTTTGT
GCTATGTGCTTATATTGTGTTTGTTAGTGCTTTATTGAGGACGTCTGCAT
CCGTGTTCATCTGGGGTACTGTCTGTAGTTTGCTTATTTTGTTGTCTTTA
CCTGCTCTGCATTTTAGAGTAATCCTGGATTTATAGAAAGCATTTGGGAG
TAGTCCTTCTGTTTATTAAAAAAAAAAATTAAGAATGATTGGTTGTTGTG
TGGTGGAATTCTGCTGTGAACCCATCTGGTTCTGGACTCTATTCGGAAGG
CTTTTTATTACTGTTTCAGTCTCCTTGTTTGCAGTGATCTATTTAGGTTG
CTAATCTCCTTATGATTCATTTGGATGAATCAAGAAATTAATCCATCTCT
TTAGATTTCCAGCTTAATGGAATATGAGTGTTAAAGTATTTCTTTATAGC
ATTCTGTATTTTTTGGCATCTGTTGTAATATTTCCCTGTTCTTTCTGTTA
ATCTCTTTCTTTCTTGTGGTTAGTTGGGCTAAGAGGCTCTTGGTTTTTTT
TTTTTTTTTTTATCTTTTTAAAGGACCAGCTCTTAGATTCATTAATTCTT
TGTATTATTTTCCTTGTTTCTTTTTCACTGATTTCATTTTAGATTTTATT
ATTTCTTGCCATCTACTGCGTTTGGGTTGGTTTTAGTTATTTTTCCAAGA
TTTTCAGTTTCATCACTAAGTCATTCATTTGGGCTCTTTTGGGTTTCTTC
ACGAGAACCCAGTTGGGACTGTTACCTTCCCTTTTAGACCTGCTTTTAAT
GTGCCCCAGAGATTTGTTACATTGTCTTTTCGATTTAACTTAGTTTCAGG
AATATTTTGATTTCTTCTTTGACCCATTCATCATTCGGTAATGAGTTGTT
TAATCTCTAGTGAGTTTATACATTTATTAGAATTTTGTTTACTGATGATT
TTAAGGTGTTTGGCTTTGTTTGTTTGTTTGTTTGTTTTTTCGAGACA
GGGTTTCTCTGTTGTAGCTCTGGCTTTCTCTATGTAGAACAGTCTGACCT

FIG. 3D(14)

CAAATTCACAGAGATCCACCTGCCTCTGCCACTGAAGTTCTGGGATTAAA
GGTGTGTGCCACCACTACCTGGCTGATTTTAAGTTTTATTACATAATAGA
CAGGTAGGGTACATAGATTTTCTACATTTGTGAAGGTTTGCGTTGTTTGT
CAGCATGTAATTCTGTGTGCTGCTGAGGGAATGTATGTTGTTTTGACAGT
TAGGTGGAAAAGTCTGTAGACATCTGTTAGATCCATTTTACATTTCAAGA
AGCCATTTAATTCTGAAGTTTCTCTGCTTATTTTTTCCCAGGTGACTTAC
CTATTGGAGAAAATAGGGTGCTAAAATCATTTACTATTATTGTTTTTTTT
AAGAAGAAAATAATTAATTTAAAAAACCCTGGAAAGAAAGATACCAAATG
TGAATCATGTTTCCTGGATAGTGGGGTTATATTTGATCATTTATTTTTCC
TCTCAAATACTGTGAGTTTTTACAATGAATAACAACATAAATATTTTTAT
GTTGCTGTGGACTTTAACTTTGCTTTGATAATATATTTGGTTTTTTGAGA
CTAATTTCTTTTTGATATTTTATTTTCTCATACTAGTTTTTAGTAAACTT
TGGTTTTGTTTTGTTTGTATTTTTTAGACTGGCCACCAACTTGCTATGTT
GTCAAGGGTGGCCTTAAAATCCACACCCAATACTTTGTCCTCTCTTTCTT
TCTTTCTTTTTTTTTTTTTATTGGAACAAAATTTCTAGGTGGGAATCTCAC
TATGTTACCCAGGCTGACCTGAAACTTCTGGGCTTAAGCAAGATGGGTGC
ACATGATCAGAGACGCTGCGCTGCCCGCCTCAGCCCCTGCTAGTTGGAAC
TATAGGCACAGACAGCTGTACTTCACTCATTTCAATGATTTAACATTTAG
ACTATATGCAAATAAATATGAAATGTATTCACCAAGTTCTCCTATGGGAG
AAACAGAGCCCTTAAGATTTTTTCCTTTCAGCTTGCCAGTGCAACGGACA
CAGCAAATGCATCAACCAGAGTATCTGTGAGAAGTGTGAGGACCTGACCA
CGGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTGACCCGACT
AATGGAGGCAAATGTCAGCGTAAGTCACACAGGTCAAGTTAGTCACAAGT
CAGGTACAATAGTACAGTACCTGCAGTTGACTTAAATATCTTAAAGGGAA
AAGGCCTCTTGGTTTGGGATATTGCCTTTCTTAATTATGTTAAATTGTTA
AAAGTTTAACTGAGGGGCTAGAAATGTGGCTCAGTTGGCTAAGAACACTG
ACTGTTCTTCTAGAGGACCGAGGTTCAATTCCCAGCACCCACATGGCAGC
TCACAAGTGCTTGTAACACCTGGGATCCAACAACCTCATACAGACATACA
TGCATGCAAAACACTAATATACATAAAATAAATCCATTAAAAAGTGTTTG
ATGATGCTGGAAGAGGAAAAAAGGCTCAACTTGTGGGTTTGGGAGCAGTT
AGTTAAAGCAACAAACCGACAGTAAAGGAGCTAAGCTTTTATTTCTTCAG
CAGAGGCATAAACAAGGGGCCGAAGTCACTGAGGCACCAGCTGCCTTTAT
TCCATTTCCCTCCCATGGAAGCACATCAGCTCAAGTCAAGCAGAGCAGCC
TGGGATGGGAGGTCATCTCATTGGAGAAGGAGGCAGGAGGCATTGTGAGG
GGAGGGAGGACAAGGCTGGGAATGGGAAGTCCTGAGCTCAGAATCAGAAT
GAGGACAAGATCTTCAGTTTCCTTCTTAATATAAAGAGGTATCACAGAGG
TCTCTATAGAAGTCTACTGGAAGCCTCACACAGGCACAAGGGTACATTTG
AAAAACTGTGACAGCCAGGGAGAGTCCCCTTCTGAAGTGTCCTTCCTCAG
AGACTGCAGCACCTGACTGTGCCCCAGTCTGCAAGAGGTTTGGGGAGAGC
AACTGACCTCCTGAGGACCCCAGATGAATCTTTAAGATGGCCTGCTTTTG
GTTTTGGTTGGTTGGTTTTTAGACAGATCTAGGAGAGTTGGTGATGAGCT
TGAATTCTCTGTCCTCCTGCCTGACCTCCAAATGCCCAGCTTCACATGGG
CTCCCATTAAGTTGTGAGTTTCGGTGTCTGGCTCCTGCTCTCACAGCCAG
TGCAGTACATTGAGCTCCATAGAGATAGCGCCGGGGCAAATGAGAGCTGG
ACGGGCACTGGGTGACTCTGTGCCTTGTGCCGGAAAATCAACTAAACATG
GGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCCTCATATGCACT
CTTTGTGAAAACCTGCTGGGAGGAGCACAAGAAGAAGCACCCGGATGCTT
CTGTCAACTTCTCAGAGTTCTCCAAGAAGTGCTCAGAGAGGTGGAAGACC
ATGTCTGCTAAAGAAAAGGGGAAATTTGAAGATATGGCAAAGGCTGACAA
GGCTCGTTATGAAAGAGAAATGAAAACCTACATCCCCTGCCCCCAAACAG

FIG. 3D (15)

GAGACCAAAACGAAGTACTAGGACCCCAATGCACCCAATGCCTTCTTCGG
CCTTCTTGTTCTGTTCTGAGTACCTCCCCAAAATCAAAGGTGAGCACCCA
GCTTATCCATTGGTGATGTTGCAAAGAAACTAGGAGAGATGTGAACAACG
CTGCAGCAGATGACAAGCAACCCTAGGAGAAGAAGGCTGCCAAGCTGAAG
GAAAAGTACGAGAAGGATATTGCTGCCTACAGAGCTAAAGGAAAACCTGA
TGCAGCAAAAAAAAAAAAAAGGGGGGTGGCCAAGGCTGAAAAGAGCAAGA
AAAAGAAGGAAGAGGAAGATGGGAGGAGTATGAGGAAGAGGAGGAAGAAG
AAAGATGAAGAAGAATATGATGATGATGAATAAGCTGGTTCTAGTTTTTT
TCTCATCTATAAAGCATTTAACCCCCCTGTATACAATTCACTCCTTTTAA
AGAAAAAAATTGAAATGTAAGCCTGTGTTAGATTTGTTTTTAAACTTTAC
AGTGTCTTTTTTTTTGTATAATTAACATACTGCCGAATATGTCTTTAGATA
GCCCTGTTCTGGTGGTATTTTCAATAGCCAGTAACCTTGCCTGGTACAGT
CTGGGGGTTGTAAATTGGCATGGAAATTTAAAGCAGGTTCTTGTTGGTGC
ACAGCATAAATTAGTTATATATGGGGACAGTAGTTTGGTTTTGGTTTTAT
TTTTGGGTTTTTTTTTTTCATCTTCAGTCGCCTCTGATGCAGCTTATATG
AATATGATTGTTGTTCTGTTAACTGAATACCACTCTGTAATTGAAAAAAA
AAAATCGTGGCTGTCTTGACATCCTGAATGTTTCTAAGTAAATACAGTTT
TGTTTTTATTAATATTGTCCTTTCGACAGGTCTGAAAGTTTTCTTCTTGA
GGGAAAGCAGTCTTTTGCTTTTGTCCCTTTTGGGTCACATGGGTTACTGC
AGTGTGTATCTTTTCATATAGTTAGCTGGAAGAAAGCTTTTGTCCACACA
CCCTGCATATTGTGGTAGGGGTAACACTTTCATCCATATTCAAAGAATCT
CCAAAATCGTGATCAGTTGGATAAGAAATATTATATAACCTACTTGGCAA
AGCAAGGTGTGATCAATTCTGTCACACCATGGGATCATTAGAATCAAGCA
ATCTGAAAATCTGTCCTTAAAGGACTGATAGAAAAGTATTTTCTAATCCT
TATACAAAGGCTCTCCTTTAACTGCCACTGCTATGTAATGACAGTTATGT
TTTGCAGTTTCCCTACTAAAGAAGACCTGAGAATGTATCCCCAAAAGCGT
GAGCCTAAACTACACAACTGCAGTACTATTTGTTGACCTTAGTCCCAGCG
AAGGCTATCACGAGAATGCTAGCTATAATATAATGCCTCTGCCCCTCTAT
CTAAATATGGATTGCTCAGGAAACTTGACTGCTTAAAGGTATTTTTTTCA
TATTGTTGTTCCTCCTATAGGGTTGCAGACCCCTTTAGCTCCTTGGGTAC
TCTCTCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATC
TATCTCTTGTCAGATTTCTTTTTTCTTTTCTCTTTCTTTCTTTCTTTTTT
TAAGATTTATTTATTATTATTTCTAAGTACACTGTAGCTGTCTTCAGATG
CACCAGAAGAGGGTGTCAGATCTCATTACGGATGGTTGTGAGCCACCATG
TGGTTGCTGGGCTTTGAACTCAGGACCTTAGGAAGAGCAGTCGGTGCTCT
TAACCACTGAGCCATCTCTACAACCCTTAAAGGTATTTTTAAGTAGTTGA
GTCAGCTTTTAAAATTATGCCAGAAGTGTCAAAAGTTCAAAAGTTTAGGA
CCATCCTCTATTGAAGTACAGGGTCATCCTGGGCTACATGAGACCCTGCC
TTAAAACCAAAATCAAACAAACAACAGGAAAAACAAGAGTTAAGAAAGAG
AAAAAGAAGCACTTGGAAACAAAGATCTGTGGAGTATGTATAGGCTTCTC
TACAACAGGTGTATGTAGGATCTTGATGGCTTTTGAGTCTATTACCCTCA
AAGAGGTACTGAGAAACCTAAATGTGATCACCGTGGTCTCTGAGGGGCAC
CTGGCAGGATTATGGGAGATAACTAAAGCTTGCTAATCACAGAGTTTAGG
GAGGGAGGACGTCTCTAAGGCAAGTTAACTGTCTGGTTTGAGATGCTTAG
GTGATGTCTGAGGAAGTAATAAGGCCTGTCCATTTTCATACACACTCAGG
CCTTAAGTCTGGGTAATGGCTACTTGAACATAAAATAGTCCTCTATGAAA
GGAATAATATCTCTGTGTCAGCAGCCTTCACGGCTAATGTTAATTGTGCA
GGAACCCTGCTTCTCAGTCAGACAGAAGCTCAATCAGGCAGGGGCAGGAC
TTCTTTGCCTTTCCCATGTCCTTGTAATTTCCCTGGCTTTTCATCTTGGT
TCAAACATACTTACCTGTTAGGTAATTATAAGAACACCAAATATTACTGA

FIG. 3D(16)

ATAAAATGTGTTTATGACTTTGTGGTGACTGCCATTCAAGAATTAGATGC
CTTAGCCAGCAATGATGGCACACGCCTTTAATCCCAGCACTTGGGAGGCA
GAGATAGGCAGATTTCTGAGTTCCAGGACAGCCAGGGCTACACAGAGAAA
CCCTGTCTCGAAAAAACAAAACAAACAAACAAAAAGATTTCGATGTCTTT
ATCACCCAAATCAAGTAACTTTCCAAAGTCTCACAGTGAGATGTAGCCTA
GTTGGGAGCCACATCTAATATATGCTGATGATCTTAACAAGTAGCCTGCT
TGTGTCTTCAGGTGACCACCCCGGTGTCCTCAGCTACCTCTAGAAAGATC
ACACTTTCCTCTGTGGTCTCTGCAGGGTCCCTGTATGATTCTGGAACCTT
GCTGTACTTCTCAGAGTCCTGATTCATAAAGCACTGAGTTTTTGCTTGTT
TGTTTGTTTTGATACTATTGGTAAGAATATATATTGAACCTTGACATGCC
TTTTTAAAATAACATTATTTTTACAATAGTACTTTAGCCTTGATTATGTT
AACTGCTTACTGTTTCAGATGACATTCGTACATCTTTTAATCCTCAAACC
AGTCCTATGAGATGGCTAGCATCATTGTCACATCATTTAGGCAAGGAAAC
AGGTCTTGGGTTAAGCTTCATGCTCAGAGCTCCTTGGAACACAGTGGACT
CAAGTGCAAGCAGACTGACGCGACTGGGTTTTACTAATTCAGTAAGCCTG
TACTCTATGGAGGAAGAGTTTCTGACCACTGGATGCAGTCTGATGACCTC
TGACTGTTCTGTTTGAAAGGTTTCTTTCAGTGATTTTATTTTTCTCCATG
TGGACTTTTTTTCCAGCTTTTAAAATATATATATATATCTTATTCGCTTC
ACATCCTGCTCACTGTCCTCCCTCCCCTGTCATCCCCTCCTACAATCCTT
CATATCCCCCCTTACCTTCTGAGCAGCTGGGAGCCCCTCTGGGTATCCCC
ACACTCGGGCACATCAAGTCTGTGAGGCTGGACGCATCTTCCCCCACTGT
GGCCAGACAAGGCAGCCCAACTAGAACATATCCCACAGACAGGCAACAGC
TTTTAGGATAGCCCCTGCTCCAGTTGTTCAGCACCCACATGAAGACCAAG
CTGCACATCTGCTACATATGTGCAGGGAGGCCTAAGTTCAGCCCATGTAT
GTTCTTTGGTTTGTGGTTCAGTCTCTGAGAACCCCAAGGATACAAGTTAT
CTGACTCTCTTAATCTTCCTATAGAGTTCCTATCTCCTCTGGGGCCCACG
ATTGGTGTCCCTATTGCTTCACTGGGATTCCTGCCTGGCTACACCCACTA
TGACCAAGGCAAGTCTTAGAAAAGACAACATTTAACTGGGGCTGGCTTAC
AGGTTCAGAGGTTCAGTTCAGTATCATCAAGGCAGGAACATGGCATCATC
CAAGCAGGCATAGTATAGAAAGAGCTGAGAGTTCTACAACTTATCTGAAG
GCTGCTAGCAGAATACCGACTTCCAGGCAGCTAGGATGGGGGTCTTCAGA
CCCACACCCACAGTTGGTGTCCCTATTGCTTCACTGGGGTTCCTGCCTGG
CTACAGGAGGTAGCCTCTTCAGGTTCCATATCCCCAATGCTGTGAGCCAC
AGTTAAGGTCACCCACTATTGATTCTAGGGTGTCTCCCTCATCCCAGGTC
TCTTTCATTGTGGAGATGCCCCCCACTTCCCCACCACTGTCAGTTGCAGA
TTTCCATTCTCGGGACCATCTGGCCATGCCTTCTGTTTCTCCTCACACCT
GATCCCGACACCCCCGCCCATTCCTTCTCCTACCTAGTTCCCTCCCTCCA
TATGCTTCCTATGACTATTTTATTCCCCCTTCTAAGTGAGATTCAAGCAT
CCTCACTTGGGCCGGCCTTCTTGTTTTGTTTCTTTGGGACTGTGGAGTGT
AGCTTGGGTATCCCATTTTTTTATGGCTAATATCTGCTTATAAGTGAGTA
CATACCATTCGTGTCCTTTTGGGATTGAGTTACCTCACTCAGGATGGTAT
TCTTAAGTTCTATTCATTTGCCTGCAAAATTCATGATGTTTTTGTTTTTA
GTAACTGAATAGTAGTCCACTGTATAGATGTACCACAGTTTCTTTATCCA
TTCTTCAGTTGAGTGAAATCTAGGTTGTTTCCAGTTTCTGGCTATTACAA
ATAAAGCTGCTATGAACATAGTGGAGCATGTGTCCTTGTGGGATGGTAGA
GCATCTTTTGGGCATATGCCCAGGAGTGATGATATAGCTGAGTCTTGAAG
TAGAACTATTCTTAGTTTTCTAAAAAACCACGAAATTGATTTCCAAAGTA
GTTGTACAAATTTGCACTCCCTCTAACCAAGCAAGTGAAAGATCTGTATG
ACAAGAACTACAAGTCCCTGAAGAAATAAACTGAAGAAGATATCAAAAGA
TGGAAAGATCTCCCATGATCGTGAATAGGTAGGATTAACAAGGTGAAACT

FIG. 3D(17)

GGACATCTTACCAAAAGCAATCTAGAGATTCAGTGCAATCCCCATCAAAA
TTCCAACACAATTTTTCTGTAGACCTTGAAAGAGCAATTCTCAGTTTCAT
ATAGGAAAACATAAAGCCCAGGAGAGCCAAAACAGTTCTGAGCCATAAAC
GAACTTGTGGAGGAATCACCATCCCTGACCTTAAAGCCGCACTACAGAGC
AGTCGTGATTAAAACAACAACAAAGGCTGCGCACTTTTGGTACAGAAACA
GACGTGCTGACCAATGGCATCCAATCCAAGATCCAGAAAGAAACCCACAC
ACTATAGTTTTTTTTTAAATATAAAGTTCTTCAGCTTAATGCTTCTCATT
ATTCATGAGAGAAGAAGACTCAACAGCAAAGAAGGTGAAACAAGGGTGAC
AAGTACCACAGGGCTCTCGAGTGTCTCTTGTGATGGACTAGGGAGCCCGT
CAGTTCTGAATGCTCAGGAATGTGGTTCACAGTGTGGCCACAGTACAGAA
GATCCCCGAGATAAGGCAGAAGACAGTCACCACAGGTCATCTCCACAGGG
CAAGGACTCAGTATATGGCATATTACTAATGCTCTTAAATATTTACTGAA
CAAAGGAACAAAATGCTGAGTCTGTCACAGAGATGAAAATAGCCGTTGCT
TCAGGGGACAGCAGAAGATAGCCTTTTTTTCTCCTTGAATGGTAGTTAAT
TTAATGTTGCCTCTATATTATTAGAAATAAATTACAAGCTGAAAAATAAT
GAGTCATACGCAGTGATTTCTCTTGCTTTAGGCTGTCTTTACTACAAACC
CATTTCAGGCTAAATGATTTTGTCTTAATCACAGTCTATGGTAATCTGTC
AAGCCAGTTGTGACCTGTCTTCCTTTCCTTCTTCCCAGCATGCAAGTGCA
ATGGGCACGCATCACTGTGCAACACCAACACCGGCAAGTGCTTCTGTACC
ACCAAAGGTGTCAAGGGGGACGAGTGCCAGCTGTGAGTACCACACACACT
CTGTGTCTCCAGTGGGGGACTGGGCCTTGCAGCTGCCTGGGCCCTGTCGG
CCACCTGCTTGCCTGGGCATTGTTGCCCTTCACTCCCAGGGTCTTTGAGT
GGACTAGTGTGGAGGTTTACCTTTTTTCCTTCAGACAGGTTATCTCAGTT
ACTTTAATATTGCTCTGATAAAACATATGACCAAGGCAACTTACAAAATA
AAGCCTTTAATTGGGCTTATGACTTAAGAGCATTGGAGTCTACATTGAGT
TCCAGGGCAATAGAGCTACATAGTAAGACTGTATCAATCAATCAATAAAT
AGGACTACATAGTAAGACTGTATCAATCAATCAGTAGATGAAGAGAAAGA
AAGAAAGAAAGAAAGAAAGAAAGAAAGGAAGGAAGGAAGGAAGGAA
GGAAGGAAGGAAGGAAGGGGAAGAACAAAACAAGCTTAGATAGGAAGAAC
AGGATAGAATGAATGACAAATGCTTGAAAAATGTTTTAGCTGTACTTTTA
GAAGCATACTCAATCCACACAGAAGTAAAAATGTTGTTCCTTATGAGTAG
TACCTAGCATTATTACATATGTACTTGCCTGTGTCCTTGGGCAAGTATTT
GTTTATTTGTTGTTTTTATACTGTTGCTGGTGTAAATTACTGAGCAGTTA
GCAGAAACATTCCTGCAAATGGGATAGTCTCTCTGATCTGAATAATGATA
TAGTTTATGTAAAAGGATTTACTTGGTTTAAAAATAAATATAGAGTCTGT
GCTTTAAATGTCAATAGAAGATAATTTCTTTTTTCCCTAGATGTGAGGTA
GAAAATCGATACCAGGGAAACCCTCTCAAAGGAACATGCTACTGTAAGTT
TTTGTAATTGTTTCTAGAGAGTAATTGAACAAAACGACATTGCTTTTTTT
TTTTACCATTGTCTGAGAATGATAAATGCTTGGGGGATGAAGCAAATACT
CATAGCCATGCCCCTGACTTGGTGAACACTGTTCTAACTGAGGCATGGTC
TCTGCTGGTCATCCAGAGCAGTTAGCAGGGGTGCTGTCCTGCCTGTCCTT
GTTCAGCTCCCGCGGAGGCGTGCTCATTCACCATTGCCCAGTGTAGCTTA
TCATGTCCAATCTTCAGACAGCCAGGAAGGAGTTTCTAAGATAGAGGTGC
GTTCCACCATTCTCTCTGCAGCTGATTTGTGCTCACAAACAAGTAAATAA
AACACCAAATTAATACCTTGGTGTGAAAGTGAATCTGGTAAGCTTACAGC
TTTATCATAAATATATTTTTTGTCTATGAGAATCTACATAGTAGGTTCTA
GACTATAGAACAATAAAAAAAGGAATTAACATTTGGCATATGCAGCATAA
TGGTATATATAAATTGTAGAAGAAAATGGATGGTTCTAGACCTGAAAAGA
CAAGAAAATTGCTTGTGTGTAATCTGGGCAGGTCTTAAGTTGTGACCTTC
AACATCTGCTTCCCAAGCAGCTGGAACCACCAGGCCTACAGAATTCTTAG

FIG. 3D (18)

CTATGATTCTAAAGGTCATTCATCAAATATAATGTTAATGTGTATTTTAT
TAAAGTTTCAAACTTCTATCTTTAATAATCTGCAAATGTAGCTCAGTAGA
GGAGAGCTCTCGCTGTAAGGTCCTGTGTTCTATCCCCAGCACAACAAAAC
AAGACATTTAAGAAAAAATTAAAACAAGTTGGCTGTATTGTCTCAGTATC
TCATCCTTGAGATAGTGAGGCAGGAGGACTTTTAGTTTGAGGCCTATGTG
GGTTATGTAGTGTGAAACCTTTCTCAAATAATATTTACACTTTTTTCTTT
AAAAAACAACTTTTTTCTTAATTTATGTGTTTTGCAACATGTAAGTCTGT
GCAATGTGAACATATCTGTTGCCTTTGAATGCCAGAGAGGGTTTCAGTTT
TCCTGGATCTGGAGTTACCAAGGGTTGTGAGCTGCCATAGTGGGTGCTGG
TAATGAACTGAGTCCTCTGGAAGAGCAGCCAGTGCTCTTAACTGCTGAGC
CATCTCTGCTGCTAGGTACTCCCCCTTCCCCCCTTAAATTTAAGACAAAG
GTCTCACTGTGTAGCCTCAGATGGTCTAGAACTCAATTTGTAGAATGGTT
GACCTTTGAACTCACAAAACTCTGCCTGCTTCTGCCTCCTGAGTGTTGAG
ATTAAAGTTGTATGTCACCACACCTGCCCCTATGATTTCTATATTTAATA
AAGATCATGACTAGGATATAGAGAACACTTTTAGAACTGAAGAAGAAGAC
AGTTACAGTTAAAAGCAAAACAAAAACAAAAACAAAACAAAACCCAGAAA
AAAAAAGAATGAAAACTAGCACTGAAGAAAAAATAAATTTTAAAAATAGG
CAAAGAGTCACTATTATATTGTGATGGATGTGTTATATGTTTAAAACCAC
AAGTGAGATACAGGCCTGAAATGACTTTAATCGAAGCTACACCAGCCTGG
GGTGGTAGTTCAGTTGGTAAAGTTCTTGCTATGCAAGCACAAGAAGCTGG
GTTTGATGCCCAGGACCCATGCTGAAACCCAGGAGTGCTGCTGAGTGCTT
CAGCTCTGGGGTGGCAGGGCTCACTGGCAGGAAGCCTAGGCTAAGAGAGA
CTCTGTCTCGAAAAACAAGGCCGATGGCACCTGATGAACGGCATCTCAGC
ATGACCTTTGCTCGGCATATAATGTGTACACACAAATTCATAGTTTAGTA
GAAGACAAGTATGATCTGCTTTTCATGAAGTCTGTTGTAATACGCCTTCT
TTAGTTAACCATAGTTGCTTAAAAAAAGAAAAAAATCGACCTCACTGGAC
AGAAAATGGATAGAGTGTTCTAATAGCCAATTCAATTCATCATCATTATC
AAAACCTATAACTTAGGGGGCTGGAGAGATGGCTCAGCGGGTAAGAGCAC
TGACTCCTCTTCTGAAGGTCCTGAGTTCAAATCCCAGCAACCAGATGGTG
GCTCACAACCATCCATAAAGAGATCTGATGCCCTCTTCTGGAGTGTCTGA
AGACAGCTACAGTGTACTTACATAAAATAAATAAATAAATCTTTAAAAAA
AAACACCTATAACTTAAACTTATCAATAACTTTAACTTTCCTACCCCATG
CTTCCTAGTTACCCATTCTGCTTTCTGTTTGTATGATCCTGGGTATGGCA
TCTTAATGGAACCACAGTGTTTGACTTTGTATCTACTTAATATTAGGCAT
GATGCCTCTGACTCTCATCCCTGATATAGCACAGTTCAAAATTGCCTTTC
TTTGGTGCTGTACATATAGCTGAGCGTTTGAGTGCTTCCCTGCATGCACA
GGTTTCTGAATTCAATCCCCAGCACAAAAAATGATAAAAAGAAAGCAAAA
AGGCTTATTTTTACAGCTGGACAGATCATCCTGCATTGTGCCTGTCATGT
TTTGCTTGTTTCTTCTGTCAGTGGACACTGTGTTACTTCTACCTTTTGGT
TGTTGTCAGGAATATTGTAAACATGAGTGAATATACACCCAGAAGTACAA
CTGGATGTGGTAATTCTATGAGTGTTTTGTTTTTTGAGGGATGGTTATTA
TTGTTTCCATACAATAAATTACATTTCCTTACAGTTCATTACATTTCCAA
AAGCCATGCATAGCATTTCTGTTGTTCTACATTCTTATTGACACCAGTTT
TCAATTTACATTTATTTTGTGAGTTTTTTAATTGGTAACCATCATAATGG
ACATAAAAAATAGCTCATTGTAGTTTTGGTATTTGTATTTCAGTAATGCT
TGGTGTGATTATCTTTTTATATTCTTATTAACCATTAGTGTGTATCTTTT
TTTGGAAAAACACCTCTTCAAGGGTTTTACTATGTAGCTCTGGCTGGCCT
GGAACTTGTGCAGACCAGGCTTGCCTCCGGTTCCCACTGTCTTAGGTAGG
TTTCCATTGCTGTGAAGAGGCACCATGACCAGAGCAACTCTTACGAAGGA
CATTTAATTGGGGCTGGCTTACAGTTTCAGAGGTTTAATCCATTATCATC

FIG. 3D(19)

ATGGCAGGAAGCATGGCAGCATCCAGGCAGATGTGGTGCTGGAGGAGCCG
AGAGAGTTCTATATCTTGATTCAAAAATAGCCAGGAAAAGACTGTCTACA
GCAGGCAACCAGGAGGAGACTGTCTTCCATATTGGGCAGAACTTGAGCAC
TAGGAGTGTTCCAAAGCCACCTACACAGTGACACAGTACATCCAAAAAGG
CCACACCTATTCCAACAAGGCCACACCTCCTAATAGTTCTACTTCTCATG
GGCCAAGCATACTCAAACCACTACATCCACCTACTTCTGTCTCCCGAATG
CTGGGATTAAAGGCATATGTTGCCATTACCCAATTTTAAACCAGATTATT
ATTGTTTTTTTGTACAACAGACTTTTAAGGTTAAAGTTTGCAGCAATAGG
CATTCTTTGAAGCTGTATCACACTGATATATGTCTGTTGTTTTCTTCCTT
CCTAGATTAAAATAGTACAGTATATTCAAGTTTCAATTGTCCCTTTCCAT
AAGAAGTCCTGGTTTCTGTTCCATTATTAGTTTATATCTTAGTGTTCTTA
AGTAAAAATACTCAGTATTTATAGATGAGTTAGATTAGAGCCAAACCCCA
ATCAGGGTATTGGTAATGAAGGTTTGCTGGATAATTCAAAGGATACTGCA
AAGATCTGGTTTCTAATGGAAAGAACATGTAAGTTGGCCATTAGTGGACC
ACACATCTGTATTTCTTATTCTTTGGAACCTTGGGCAGGATAGACAGATG
AGCTAAGATTCCTTCATAGCTATTGAATTTGTGAGAAAAACAAATTGTGT
TTCCAGAAACCTGCTTTAGTTTGTATCAACACTTACTTTCTTTCTGTGTG
TGGTGTGTGTGATGTGCCTGTACCATTTTCAAGTTTTTCTTCCTTCTTTC
CATAGATACCCTTCTCATTGACTATCAGTTCACCTTTAGCCTGTCCCAGG
AAGACGACCGCTACTACACAGCCATCAACTTTGTGGCTACTCCTGATGAA
GTAAGCTTTTCTTTTAAGCTGTCTTATTTTGTGTTAAATTTTGTATAGGT
TTTTTTCTTGGTCATCCTGGACAAAAGTACTACATAGAAGCAGACAGTAT
CAGGGTGGGAATATAAAAGGCAACCAGTTTTTAAGTATTTTTTTATTTAC
TTGTTGACAGTTTTATATGATTATATAATGTGCTTGATGATATTCAACCT
GTGACCTTTTGTCTCCCTCATACTTAGTTCCTTCTCTCCCCACCAAGTCA
CCTTCACTCCCTCTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG
TGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA
AGACAGACAGACAGATAGACAGAGAGACAGAGATTGATTGATTGATTGAT
TGATTGATTGATTGATTGATTTACCTACCTAGTTTACCAGCTGACTGCAG
GAGCATGCTGGGTGGGAAGTTCTTACTGGAGCATAGACACATTACAGTGA
CTACACCACTGAAGAAAGTGACTCCCTCTCAGGTAGTCTTCACTGCCACT
AGGTCCTCAGGGATCAAGAGAATGTTTGGAGTCTACATTTTATCTTTTTT
CCACTCAGAAGGCAAACATTACTGAATGTTTTTAAGTAGTAGAATAATGT
TCATGATAGTCTGTTTAATATTAAATTAAGAATTTGTTCCTAATTATAAA
ATTTTTAGAAGATAGACAAGAAGACAAAATTTTTGAGTTAACAGTTTGAA
AGGTTTATTTTTATTTTATTTTATATGTATGAATATTTTAGCTTCTTGTA
TCCCTGTGCATCATGTGTGTGCAGTGCCTGTGGAGGCCAGAAATAGATAT
TGGATCCCTGGAACTAGAGTGATAGATCATTGTGAGCCATCATATGGGTG
CTAGAACCAACCCAGGGTCCTCTGCAAGAGCAGTGAGTGCTCTTAACTGC
TAGGCCATTTCTTTAGCCCCTAAATGTGAACAACTCTTTAAATAAATGTA
AGTGATCTTAAATACTCTGGAGAAAAATCTGTAGCTATACCTTACTTTTT
AAAAATTATTTTGTTTTATATTATGAGTGTTTTGCCTACATATATGTGTG
TCTGATGCCTGCAGAGGTCAGAAGAGGGTGTTGGATCCCCTAGAACTGGG
GTTACAGATGGCTGTGAGCAGCTATGTGGTGCCTGGGAGTTGAACCCTGG
TTCTCTGTTAGGGCAACAACTGCTTTTAACCATCAACCCATCTCTTTGGC
ACATGGGTGCATTGTTGGTTTGGCTGCTTGAGTTGTGTGTGAGGGGTGTG
TGTGCATACATATGTGGGTCCATGCTTATCCAGTGGAGGCCAGAGGTCAG
AGTCATGTATCTCTCTGTTACTTTCTACCTTATGTTTTGGAAGCAAGATT
AGATAGACCCCTGGGACCTTCCTGTCTTCTCCTCAGCACTAGGACTACAA
GTCCACACCTGACTTTTTACATGGGGCTTCAGATCTAACTCAGTCCCAAC

FIG. 3D (20)

ACTTGTTTCATTTCCTTAGCACCTTGGCTAGATTCTTAGGATTTTAGAAG
GAGCTTATAGCAAAATACCACAAGTGAAATTTACTACTGCCTTAGTCATA
AGCAAATATTGAAGGCTCAGTCTTTAAGGGTATAATTGATAGTGTTCTTT
TTTTTTTTAAGTAAACAAATAGCCTGTCATGGTAACTATCGCTGTAGTCC
CATTACTTGTGAGAGATGTCAGCTCAAGGCCAGCCTCCGCTACATAAGTA
AGGGAAGACCAGCCTGAGCTATATGGGACTCTATCAAAACAAATAAACAT
TGTAGAATTTTTGTAATACTTATTAGAAGGTAGCTGATGATCATGAGAGT
CTTTAGACATTTCTTCATTCCACTGTTTTGTGTGTGTGTGTTTCATGACA
GATTTCTTACTAGATTTATCTCTTTGTGTGTGTGTGTGTGTGTGTGTGTT
TTACAAAATGACAAAGATTTTAGTCCTTCTCGTGGAAAGTAGTTGCTAGT
GGTCAGCAGATACTTGCTAGTATAAATAAATGAGCATAGATCTGCGCTTG
CAAAGGAAGACAAAGGGAAAAAAGGTTTTCTTGAACATAATTCCTACTTT
GTGAAAGAAACTTCTCATTTGGAAATTACATTTTGAAAATAGGTATTGTG
AATGTTTCCATTGTGGTTTGTGGTATAACTATCAAATAACACTTTTTTAA
AAAGAAAAATCTTAATTTTCTAAGATTTTTAAATACCCTTTTAAAATGAG
CATTTCCAGCATGGTTTGATTAATTTGTAAAATGTAAGAATATAGTATCT
AAGGCTACAGAAATGACTCAGTGGTTAAGAGCACTGGCTGCTTTACAGAG
GACCCAGGTTCCATCCCCAGCACCCTCATGACAGTTCACAGCCATCTGTA
TTTCTAGTTCCAGGGCATCTGATGCCCTTCTCTGATTTTCTCCAGTACTA
GTGACACACAGCATACATTTGAACAAAACCACTGATACACATAAAATAAA
TTGTTTTCAAGAAACAATATAGCATCTAATTAGCTTACAAAACTAATTAT
TTGTTTCTGTACTAATTACGTTTCTATTGGCATGACTAAGGCAACTTATA
AGAGAAAGCATTTAATTTGGGGTTCACACTTCTAGTGCCTTAGATTCTAT
GAGCATCATGGTAGGGAGTGTGGCAGTAGGCAGGCAGGCATGGTGCTGGA
GCAGAGGCTGAGAGCTCACATTTGATTTTCTACTAGAAGACACAGAGAGA
GCTAACTGGAAAAGGCATGGGCTTTTCAAACCTCAAAGCCCCCCTCTAGG
AACACACACCTCCACCAAGGCCATACCTCCTAATCAAACAGTCCTACCAA
CTGAGGACTAACCATTCAGAGATAGATGAGTCTATGGAGGCCATTGTCAT
CCAAACCACCACAGGCCCCAAGAAAGATTTGTTAGTGAAATTTCAGTGAA
AACTAAAACAGCATTAGAATTTACCTGGCATAGCCAGCAATGATCTCTTC
TGTTCAGTGCCACAGATTTCTTTGAGTTAAAACTCAGTTGTTAAAACCAA
AAATCAAAATGTAATTGGCACTTTAAATTGCTATAAGGGGAAACAAGGTT
TTCAAAGCCATGAAACCATATTCAGAATAATTTTAGCGAGAGAAATATTT
TTTCTTTTTTTTGTCGTTTCTTTTTTTTCTTGGAGAGAAATATTTTTATT
ATTTTATATTATTTTAATTACATATTTAATTATTAACCATTTCTGACAGA
GGGCAAAAGGTGAGGATCTTCATGGAACTAATATCTGATAAAGCACCAAA
TTCTTCCCAACTCTGGGATGCAAATGACAGTTCAACTTCAGTTTATTGCT
TGTATTGAAGAAAATTGACAAGAAATGTCATGTCTTAACATAAGCATGGA
TTTCTTTTAAGATGTAGAATAGTCTATAATTAATGTTTTTGAGACTAGTA
AGACCTGATTATTGTTGTATCTTAAAATCTAGAAGGTACTAACAATTTTC
TAATGTGTATTTTTTTTTTCATCAGCAAAACAGGGATTTGGACATGTTCA
TCAATGCCTCCAAAAACTTCAACCTCAACATCACCTGGGCCACCAGCTTC
CCAGGTACAGACACACCTAGAGAGATGGATTGGCAAGTTTAGTGTAGGAG
TTGGGGAAGGAGGCTCTGAAGGCTGGTGAGTGAGTTCAGAGCCCACCTCT
GCCTCTTAGTAGCCATGGCACCTTGAACAAGCCATGCTTGAACAAGCATG
TACAATTCCCTCTCTACCTTAGGCTACTCAGAGTGAGGAGTCACAGCTCT
TGCCTCCAGCGTTGCTGGTTCAGGTTGGTTGGATGGCTGCTCCCTGCTTT
GCCACCACCTTCCAGCACTATGACTATCTCTATGTTTGTGCTTCACAGGG
GAAAAACTAAAGTGACTCATAGTTTTAAGAAATGAAAACTCTTTAAGGGA
AGGGGGATAACTCTAATATGTAGAGGTATTCATACTTTGGGATAACTCCT

FIG. 3D (21)

AAAAGTACAGCTTTTCCATTCTTGTTTATCTTATAGTGACTATAAAATTC
TGATGGCCCTAATGTAGCAGTTACTATAAATAACCACTCCATAACTTGAT
AGCCCTGAAGATAGACCTAGGTTTGAATTTACCTGCACGGTGTTGAACAA
GTTACTGAAGCTTTCTTTTCTTTGTTTTTTAAGTTTGTTTTATTTTATGT
GTGTGTTTGCCTTTGCCTGTATGTGTATAAGTGTACCATGTATGTGCAGT
GCTTGAGAAGGTCAGAAGAGGACATCAGCTCCCCACCCTCAACGAGTTAC
AGACAATTATGAACTACTATATCTGTGCTGGCAACAGAACCCAGGTCTTC
TGAAAGAGCAACCAGTGCTCTTAACTGCTGAGCCATCTCTCTCTAGCCCC
CAAGTTACCTAAACTTTCTGATCCAGTTTCCTTCTTTATAAAATGATACA
GTGAAAATAGCTTTGCTATGTACAGAGATATTCCAACTTTTTAATATTAC
AACATGACATCTACAAATATGTTAGCCCTCATTCATAATCTTGCCTGAAT
TGTAGAGTGTTGCAAGGAATAAATGAAATAAAGGAGGTACTTATTATAGA
GTTTGAGGTTTGCCTTCATGCATAAAGAGAAGCTTTTTTGAGTCTGTACT
ACTCATGTTCTTAGCCAATGGAGTATATAAAATATGGTAGAACCATTTAG
AAATGGAGTCTCACTGGGTACAGGCCTGAATGCAGTGGTAGCAGGTAGCA
GAAAGAAGGCCTGAGTGGCTGCTTGAGCACCTTCTCCATCAAGACTTGAG
GACCTTTCTGCTTAGGAAGTGATGAGCGAGTAAGTGTCCCTGAACAGGAG
CCTTGAGCATATTCTACAGTGTGAAGCAGAAATACAAAGGAGTTGAGGTA
TCATGTGCAAAATGAATGCAGTGTCTGTTTTATATGTATGATTGTTTTAC
ATACATGTATGTCTGTGCATCGCTTATATATCTGGAGCCTCTGAGACAGA
TTACTTAATCTATTGGGACTTGAGTTTTTCCAATCTGTAGATGGAGATAG
GAAGGTGTTGTGTGGGTTAGAGACTGAAGCTCATAAGGCTATATTCTTTT
GACACTGTAAGTGCTCAATAAACTTTTACCCTCATTACTAGTGCGCAAAG
ATTCTTTCTGATTGGCATACCCGCCTCCCAAGTCTTTATTTTTATTCTTG
CTTCTTTCTAGCCGGAACCCAGACTGGAGAAGAGGTGCCTGTTGTTTCAA
AAACCAACATCAAGGAATACAAAGATAGCTTCTCTAATGAGAAATTTGAT
TTTCGCAACCATCCAAACATCACTTTCTTTGTTTATGTCAGTAATTTCAC
TTGGCCCATCAAAATTCAGGTAAGAACTGCTTTTTAACTTCATTCCCGTA
AAGATGGTGACATCTCTTTAGTGGAGACTAACTTCACTCATTTGGAATCT
GTGGTGACTGAAAGATAGTGTTGCTTTGCCTTTGAGGGATCTTTGCCATA
GACTGAGTAGCAGGTGAGTGCTGTTCTTAGGTTGGAGAGATGTTCAGTGA
GTGGAGTGCTTGCTACACAAGCCTGAGGACATGCAGTTCATCTGCAGCCT
CTCATACAAAGCGGGACACGCAGGGTGTGCCTGTCACCTCAGCACTGGAC
ATGCAGTGTGTGCCTGTCACCCCAGCACAGGACACGCAGGGTGTGCCTGT
CACCTCAGCACTGGACATGCAGTGTGTGCCTGTCACCCCAGCACAGGACA
CGCAGTGTGTGCCTGTCACCCCAGCACTGGACACGCAGTGTGTGCCTGTC
ACCCCAGCACTGGGAAGCAGGGGACAGAAAGATCTTGCTTGCTGGCCAGC
CACTCAAAGCTGGATCTGTGAGTTCTAGATTCAGTTAGAGACCCTGTCTC
AAGTAAAATAAGGTAGAGAGGAATTGAGGAAGACACCTGATTACCTCTGG
CTTCTGTATGCATGTGCACATATATATACCTTCACACATATACACACTCA
GAGAAAAAATTCTGAGAGTGTCATATCACTTGTGAAGAAAGTTTTAAAGC
ACTTTTAAAAGCAAGATGAAAGCTATGCAAGGTATGCAAGGTAGTATACT
TTTGTAATCCCAGGATGTGGAAGACCAATGCAGGAGGATCACCCTGAGTT
TGAGGCCATAGGAAGACCCTGCCTCAAAAGGAGGGAAGGAGGGAGGGAGG
GAGAGAGAGAAAGAGAAAGAGAAAGAGAGAGAGAAAGAGAAAGAGAAAAA
GAAAGAAAGAAAGAAAGAAGGAAGGAAGGAGAAAGAAAATCAAATTGATT
GGCATATAGTTATGTGTTTATTTTTTGAGTAATTGCTATGTAAAAGCCTT
TAGAAATACACAGTTTTAATTATGGAATTGAGTATAAATAAAACAAGTAC
ATGTTTGTAACCAATAAAGTATAAAAATGACACATAAGATGTCAAAGTGG
TATGATGGCTATAATGTGGAGTCCATAGAGGAAGCAGTAGGCAGTATGAG

FIG. 3D (22)

GTACTGTGTAAAAACACATAGCTTTACTATTGCACAGACAAGTGTGGATT
CTTGTTCTGTGTGTGGTTCATGGAGGCTCTCCAGTTTGCAGATTCTCTGT
GCATGTGTCCTGAAGGATTGGTCTTCCTGCTATGACCTCTGGTGTTATTA
GCCTGAACTGAGTCCTAAGGAGACAGGTAGTGGAAATGTTTGTATTGCAA
AGACAGTATGGGTAGTTGTTTTTAGAAACAGGAGTTCAACAGAATTGATA
GAACTTGTGATCAAGAAGCTAACAGCTGGACTGGGATGTAGCTCAGTTGA
AAGAACGCTTGTCTAACATTAAGAAGCCCTGGGTACCATCACTACCACAG
CATAAACTGAGAGTAGTGACAGACTCATGTGTCCCAGCACTGGGAAGGTA
GAGGTAGGAGGATCAGAGGCTGCCCAGGGAGGTTGAGAGTGACTTACGCT
AGGAGATAGATCTAAAAATGAAAAGGAAAAAGAACTTGGTAGCTGCTAGA
GCTACCATGAAGAGAGTGGAGCTTAAGGATTCAGCTGAAGAATGTAAACT
GCCTTCTGATGACAACTGAGAGTCGCTGAGTTATTTAAAGTCAGGAAGTG
AACAAAGATCAGTGTTTCAGAAAGACCTCTGTGGCAACAGTATTGACTAG
AAGTAGCCCCTCCTATGTCAGGTACTGGTTTAGACTGTATTTGGAAGTGT
CCTCTTTCTTGATGGCCCTCAGACACCTTTCATGGCCACTCCTCTGCATT
TGTACCCCATAGCCACACACTTGATGGTTCTTTATTACATAAATAGCTCC
TTATAGGCAATGATAGATTTTATATTTTTGATAATTTTAAGATAAACTCT
ATGTCATTGCATAGAATTTAGTAGTTGTAGGTACTCAGTAAATGTATATA
GGATGAATACAAAAGCTTTAGGGTAACAGTATTTTGTTGTTCTTCCCCCG
CATTTTTAACTATCTCATAGTAGCACAGACTAACCCATAACTGACCATGA
AGCCAAGGATGACCTTGAACTCCTGTACCTTCTACCTCTTCCCCGAAAGT
GCTGAAGTTACTGGCATGTGCTGCTCACCCAACTAATAGCAAGTTTTTCT
TATAAAGGTGCTGATGCCCTTTCCCTGTTTGTGTTAATTGCTGACACTTA
AAAGCTCTTTATCCCAACCCACAGTGTTAAAGAGTTTAGTTAAATTTTGT
GGAAATTTTGTCCCAAATGAAGTGGTTGATGGCAGGCCTGGTGGCTCCTT
CCTATAATTCCAACACTCAGGAGACAGAGTCAGGACGATGGCCAAGAATT
CAAGGCCTTGGGCCTACAGAGTAGAAGAGAGAAGAATGAGGATTCGAACA
CCTGATTAAATAGATACCATTTCCTGCTACCAACCTGTGCCTTAGCTACT
CTTCTATTGCCGTGACAAAACATCATACCCAAGGCAGCTTATAAAAGAAA
GCATTTATTAGGACTCACAGTTTCAAGGGTTATACTCCAAAACCATCATG
GCCGGGAGCAGGCAGCAGGCAGGAACATCTGCTGTGAGGAAGAGCTGAGA
GCTCACTTCTTTATCCACAAATAGGAGGCAGAGAGAAAGCTAACTAGGAA
TAGAATGAGCTTTGCAGACCTCAAAGCCCACCTCCTTCCCAAACATTTCC
ACCAATTGGGAACTAAGTATTCTAATCTGTGAGCCTCTGGAGGCCCATTC
TTATTTAAACTACCACACTTTATAAGTTAATACTACATGTGATGAGGAAA
CTGGTATGGGAATTCTGAAAAGTAGTTCACAGGAGTGGGAGGGGCTGAAC
GTGAGTAGATGCTAGCATGTGTGTCAGGAGTGAAGTGTTCAGAGCATTGC
CTGGTTTGACTTCTCTCCAGAGCTGAGGTGAACATGCTTTGTGCCAATAC
AAACCCGTATTAAAGCGGTGGTAGTTACTGAAAATCAGTGCAGGGCTGTG
GTCTCAACACAATGTTTGAAAAAGAAAACAGGGCATCCACATCAGGCAGT
GTACAGCTGCTTATAATTCCAGTCCTCTGGCCTCTGCTCACATGCACATA
CCCCCCCATACATACACACATGATTAAACATAATGAAAAATTAAAAATTA
ATGCTATAAAAATGGAAAGAGCCGGGCGTGGTGGTGCATGCCTTTAATCC
CAGCACTTGGGAGGCAGAGGCAGGCGGATTTCTGAGTTCGAGGCCAGCCT
GATCTACAGAGTGAGTTCCAGTACAGCTAGGGCTACACAGAGAAACCCTG
TCTCGAAAAACAAAAACAAAAACAAAAACAAAAAAAAAAGTGGAAAGAAA
GGTTCACTGTTTCACAGGAAAACTCTGAGAGGTGATAATCCAATCCCAGT
TTAAAATATACTCCATAGTGCACACAGCCTCTCCCATCCTTGGCAACTGA
GGCCTGTGAGAAGACTCAGTCCTCTCCTGGCTTCCAACCTTACAGTGTTC
AAAACTCTTCTGCAAGATCCACATGGTCCTACCAAGACCCTGAAGGTCAG

FIG.3D(23)

GCATGCTGATTAGGCTGTCTCTGGGCCTGAAGTGAAAGGTAAACACTTCC
GAGATCTCCAAAGCCTTGGGAAGATTCTGAAATGTATGGGTGTTGGTTCA
GGTAGACTCTCAGCCTTGGTGAAGCTGCCCCCGGAGCTGTAGGGTTATCT
GCAGAAAGTCAGCCAGGTGCACTTACCCTGGAATCCTCTCCCATTCACAG
ACACCTCCCTGAGGCTTTGTGGCTTCACCTCACTGTGCAGCTAGCTCCTG
TTTTACATGCTTATATAATGAATGGTCTTGGTAAAGAAGATGATAAAGGC
AAGCTAGAGGCCTTTTTTTTCCCCTCTTCAAATTTTGATTGGCCTTTCCC
TACTGTTACACTGTCTACTCAAGGTTTTGAGCATTTACTTTGTGTACATA
GTAAAAGCAAAGTACATATTTTTAAGTAGAAAAGAAAGCATCTGTGGTCT
TTGATATAGGTGCTTTTCTTTATTTTAATAGTAATACTTATTCCATGCTT
GTTAAGAAATTCATTCACAGCGTGTTTTCATAGAGACTTTCTCTATAGAG
ATATATAGAAATCTAGACATGAGGACAGCCCACTAACCCACTCTTCAGAC
ACTAGCTGCTTCTCTTAGAGCCCTGGGCTCTCACCCTTTGGAGGACAGCC
ATCCTCACTCATATGTGACAAGCTTAGACACAGAATAATCACAGAGACTC
CAGCCTCCCCCACAAACCCACAATGCCAATATCCCATATTCCCAGGAACT
TTTAATAAGCCATCCACTCTAATACTCCATCTCTTATCTCAGGCATAGGC
CCTGGTTTTGGTTTGCTTCAGAGTACTGCCTTTTCTCTACCACGCCCTTC
CCACTCTTTGCTGACCCTCCAGAGATGTCATTTCCAAATGAAGGGGGTTT
TTGGTTCTGTGGGTGTTTTGTTTTTCAGTGCAGTTCCTTAACTGCTATTC
AGGGGACGGAGCAGGCAAACCAGATCTCTAACTTCTGAGGCCTGTGAAGA
GAAGCATCAGAACCTCCCAGGGGAGCTGTAGGAGCAGGAGTCAGGCCTAG
ATATGACTGTGAGAGAGTGGGGACCATTACCAGTGTCTTACAAATGAGGG
GAAGGACTACCGTGCTGGGCCCTGAAAGATAAGGAGGACCAGGCTTCAGG
AAGGTAGGACACATTCTGCTGACTGTCTGGGATTGAGGACAGTAACACAA
CTACTTAGACATACTTTGAATGAAGGACAGACTTAGTGCTTCAGAACTGT
AAATCCATTATATCTTTCCCAAGTCTTAGGCTAGCCAAGTTTCTCAACAT
TTATCTACCTCATCCCAAAGGGTTCCCAGGACAAATATTTCTTACTCAAA
CATTTGATGGGAGTTGGAATCAGGTTGAGGAAATGCAGGGGTGTAGATTT
TAGATTTCTGGGAATATGTATAGATAGCTACCTTCTGTTGGATAGAAAAT
GAGATTGTAAGTTTTTCAGTGTTTTTTTACACGAGTTTGTGTGCCCATGT
ATGCACATGTGGAGGCCACGGGTCTACCTTAGGTGTCTTCTTCAGGAACC
AGCCATCTTATTTTTAAGATGATCTCTCTCCAGACCTCAGGGCTATCAAC
ACACCTCAGGGATCCATCCTCCTGACTGTATGTCCCTAGCATTTGGGTTA
CTGTACCACCATGCTCAGGTCTTTGTGTAGGTCCTGGGGATCACAGTTAG
GTTCTCATACTGCAGGGCAAGCACTTTGTAAACAACTATCTCCCCTGCAT
ATGGAAGTATTACCACTAAATTACAACAAGATTTTCTTCTATTAAAATTA
TATTTTAGAAGCTGGATATAGTAATGCGTTGGGGCAAAAGGAGGGAGGGA
AATGAAGAGGATAGGAAGAGGGGGAGGGAGAAGGGAAAGAGTGGAGGCGG
GATCAGAAGTCCAATGTTATTCAAGGGCAGCCTGACCTAGATAAATCCCT
ATTAAAAAGTTTTCAGTATAGAAACTTCTCATCACCTTCATTATCAGAAA
AGCCCCTAAATTCAGAACACTTTTTAATCTTAATTAGTTGACAATTTCAT
AAATGTATTATTTATATATATGAATAACATTTTCCTCCTACCTTTTTTTC
CCTTCCCCTCTGATGATTCCCATCCTCCCAACCAAGCCCCCCTTCTGCAT
TTGTTTGTTGCTTTAATGACCCACTGAGTTCCATTGGGCTCACTTCCATG
AGTGTGACTAGAAGAGCTATTTATCAGAATGTGGGCAACTTACCAGTAGT
GACACTGATGAAGAAAGTGTTTCCCTCTTACCCAGTAACCATTAATGGCC
AGGAGCTCCTGGGAGGGGTGGGCGCCTTATGAGCCCCTTCTCCAAAATGC
TTTCAAACTGTGACCAGCTATATTTAATGTTTTTATTATGCCTGTGTATC
CATGTGGGACAAGAAAGCTTGAGAGTATCATAGCATGCATGTGGAGGTCA
AAGAACAACTGTGTAAAGTCAGATCTCACTTCCCACCTTCACATGGGCTC

FIG. 3D(24)

TGGCACTGAACTCATGTCAGTGACCTGAGAGGCACTTTATCCTCTAACAC
GCACCCTGTGCCCAGCCTAAAATTTGACCTTTGCAAGGTTTAGTGTGTGT
TATCTGACTGTCTGAGTAAGGATGACAAAATGAAACCAAACTTATGGGAT
AAAGCTTGGTGGTTGTATCAGTACATTTTTATTGTTGTGATAAAACATTA
TGACCAAGACAGCTTATAGAAGAGTTTATTTGGGTGTATAGTTCCAGAGA
GGTAAGAGTCTGTCCTGACAAGGAAGCTGTGGCAGCAAGTGGCAGGTATG
GCTACAGGAGCAGGAAGCAGAAAGAGCAAACTAGAAACAGTTGAGGTTTT
TTAATAGGAAAGCCCACTCCCCTAATGATGTCCTTCCCCTAGCAGACCAC
AAGTCCTAACCCTCCCTACACAGCACCACCAGCTGGGGAGTTCAAATGTC
TGGGACTGCAGGGGACATCTCATTCAGACCACCTCAGTGGGAGAATGCTT
GCCTTCATAGTATGTGCAAGGCCCTAGGTTCAATTCTAGCCAAGAAAAGA
GAACATGAGGAAAGAAAAGAAGGTGGGAGAGAGTAGAGAAAGAAGAGAAG
AAGAGGAAAAAGGAAGGGAAGGGGGAGACAGAGGAAAGCAGGGAAGCAGA
GGAGAGGAGAAGAGAAAGAAAAGATTAACCAGCCTGGTTTTTAATAGCAC
CCCTCCCACTCTCAGTAGTTCCCAATTTGAGCATTAAGTTCAAGACTGAT
AGATATTTCTGGGTGGGTGACCAGTGTGGTCATAAACATGGTGACTTTTG
CTCTCCGTACAACTTGTGATTATGAACTTGTTAGATGATCAGCTTCAACA
GGAGAGGGCCTCCTTTAGTCTCAGGTGCCCCCTCCAGCCACCCTGGGACT
CGCAGCCTCTCTGTGATGAGACACAGGACATTAACTGGTATGGTTCTGCT
TTGCCAAAACGTCAGTCCATGGTTGAACTCTCCACAATGAGAAAGAAGCT
TTGAGAATCATTACATGGCATCAGGCAAGCCAGGACTGATGGAGCCTGAG
AAAGGGCCAGGAGCATCCGCAGGTTTTGGCACCCAGTACTAACTAGTAAA
AGCACCTCATAGGTTTCTTTAAAATGCAAACACTAAGGAAAATCTAACTT
TTTTTTATTTATTAAGGCCATTCATTTTATTTTATAAGTATTTTGCCTGT
ATACATATGTACCACATGCATACAAGGTCAAAAGATAGTATTGGGTCTTC
GAACTGGAGGTACAGATGATTGTGAGCTGCCATGTGGATCCTCGAAATTG
AACCTAGGTCGTCTACAAGAGCAGGAAGTGCTCTTAACTTCTGAGCCATC
TCTCCAGCTCCAGAAAAGCTACTCATAAAAGTCAAATCTAAGCCATGTGT
CTGGTGATGTACACCTTTAATTGTAGCACATGGAAGGCGGAAGTAGGCGG
ATTGTTATTCATCCAAGGCCAGTCTTCTCTTAACAGTGACAAAAACAAAA
CCAAACCCGAAACCTGTTACTTTGCACTTTAGAGTATAAGTGATAGAGAA
AAGACACAGAAATTTTAGAATCTATACCTTAAAATACCTTATGGCTTATA
TGATACTGTTGGGACCATATTTACTTATGGAATGCAAAAAAAAAAAAAAA
AAAAAAAGATGGGGGGGGAGCTGAAGGTCTCCTTTCTATTCTGTTGTAAA
TCTAGCTATAAAAAGAGTAAGAGGCATGAGTGTGTCTCAGTGGTAGAGCA
CCTGCTTAGCTTGTGTGGGATTGAATGATCCTCAGCACCACAGAAGAAGG
GTGGGGCAATAAATTTAGGAAAATAAGATGCTAATCATTGACTTTCTTGA
TTTTTTTAAAAAAAGTTATTATTTTATGTTTATTGTATATGTTTATATTT
TCTATGTGTGTTTTTGATGTGTGCTGGAGGGATGGGGGCCACTTGCTGAA
CTTCCCAATTGTTATCATAACTACCATCTTTAGTGAAACAGTTACCATCT
ACTTAGTAATTGTTTCATTCGAATAGATACTGAACACTCTTAATCTGAAA
CTAATGCTCAGAAAGTTCCACTTTGCCAAGCAAGCAGGATAATGTAAGCC
TATAATTTTAGCACTGGGAGGGTGAGGCAGAATTGTGAGCTCAAGGGCAC
CCTGAGCTTTTGAGATCCTGTCTCAAATAAAATTAAATTATATAGATATC
AGATTTTCAGAATAGGTGTGTTCAGCTGCTGAATAAATCTAAGCAAATAT
CCCCAAAGAACCCTGAAATCTGAAACGTATTAGTTCTAAGCCCTATGTTG
TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG
TAAGATAGATTCTCACTATGTAACCCTAGCTTGCCTGGATCTTGCTATAT
AGAGAGACCAGGCATATGCTATCGTGCCTGGGAGTCCCAAACGTTTTAGA
TGAAAGATTTCAGTTGTACCATTATCTTCCTAATGAGGGCTCTGGTCTAG

FIG. 3D (25)

TGAGGCAGGTGACATTAGGCCAGTAGTAAGTATTAGGAATTGGTGATGAC
GGTCAATTCTGAGACACACAGTAGATACATCTAATCTACCAATACAACCA
ATGATTTAGAAAGAATTAGGCCATAGTTAAATTTGCAGTGTTTTTCTTCT
CCACAAAATAATGTTACTTCTTTCAGTTCTTAGTTCAAATACAGTAGGAA
TTTTTTATATTCTTGGTGCTAAACACTATTATTTTATAGTAAAGTTAGTA
AGATAGAAATGACGCCCTGTGGGTTGTCTGGTCGTAGTCTGTAGCTGAGG
CCATTTTGCTGAGAAGCAGCGTAGGCTGTCACTGGCTTTGTCACCCATAT
TTTCTGTATTTTTGCTGCAGATTGCCTTCTCCCAGCACAGCAACTTCATG
GACCTGGTACAGTTCTTCGTGACTTTCTTCAGGTAATTTCTCTATGCTAA
TTGTACACATTCCATCGAGACAGTCCCTTAACTGCAGCTTGCTTTGTATA
TCCCTACAAAGCTGCTTTTCACTCACAGTGATGTAAATTTAGTCTGATGT
GATAAAACTCTCCGTTTGTATGATTCGGCTCTTTGCATGGGGAGAGGTTT
GGGCTCAAGCAGTTATTAATAATATAGCTACTGCTGTGAGCTACATGTCT
TAATCTGTCTTAATCAAGATATGACTGTGATTTTCCATAGGGAAAGGTAA
GGTTTACTTGCAAACTCCTGGGGTTCTCCTTTTTTTATAGTTTTCTTATT
AGTAGGGTTTTTTTTTTTTTTTGAGAATACTATGCAGAAATGATTGAAAAG
AACAAATTAGTCATTGCATATTGGTAAGAGAAGCAGCAAGAGCCACCTCA
CCTCCCTCTGCTCTCCCCAAATAGAAACTGCTCTGCTGTGCTGCTTCTCT
ACCTTCACACCAATGCTCGGCCTGCCAACTCAGTTATCTTTCCTTTCCTT
TTAAGATAGGGTCTCTCCTTATAGTAGTTATGACTGTCCTGGAATTCTAA
ATAGAAGAGGTTGGCTTTCAAATCACAGATCCTCCTGCCTCTGCCTTCTG
AGTACTGGAAGTATGGTGTATGCCACCGTGCCACAGCTAACTCAGTTATT
TTTTGGTGTTCTATAACTGCCTTACATACATACAGACCAGGTACACACAA
AATTCCTTTCCATTAATTTAATAGTTATATCACAATGCATTGACCAACTA
AAAAATCCTAAATTGACTTATGATTCTACTTGCTCATGTTTTAAAGGAAA
GGTTACTCTTTGCTTATCTTAAATGTAATATTTTTCCTTTGCAGTTGCTG
TTTAAATTTTCCCTATAAGTCGACCCCAAATTTACATCTATAATCTGGCA
AAACAAAAAGACCTCTAGTGATGGTTGTCTCTTAGCTTTAGTCTCTCTTG
GACTCCATTCCCTCCACCCATAATGTTCCATCCTCTGTCCTTAAGTGTAC
TAGTCTCCAAGGCCTGCTATGTGGTTGTCATTGTTGTAGTTACTTTTCTA
TGTTGTGACAAAGCACCCTGACAGTGGCAATTTAGAAAGCATATAATTTG
AGGATCACAGTTCCTGGTTAGAATCCATGACCATCTTAGCAAAGGCAGAC
AGGCAGGCCTGGCACTGAACAAGTAGCTGAGATCGTCCATCTGGTCCACA
AGCATAAGGCAGAGAAGCTAATTGGGAATGGCATGGGCTTTGGAAACCTC
AGAGTCCACTCTTAGTGATACCTCCTTATCCTTCCAAACAGTATTACACA
TTCAAACTTCAAATGTGTGAGCCTCTGGGGACCACTCTCATTTAAACCAC
CACAGTGATCTTGGCAACTTCTTTTGTGTTCGTCCCATGCCACAGTCTTT
CCATGTATTTCTCCTTTTGCTGGAACTTTTTCCCTCGAAGGTTCCTGAGG
AAAGAAACATAGATAACTTTTGTATGTACTTCTACAACTGAAAGTATCTT
AATTTTTGCCCTAACAAATTTTTGTTTGCTTACTTGCTTGCTTACTTGAT
TCTGCGTGCATGCATTTATTTGTTTGTTTGTTTGTTTGTTTGTTTGAGAC
AAGATCTCTCTTTGTAGTTCTGGCTGCCTCAAACTCAGAGAGATTCATCT
GCCTCTGCCTCCAGAATGCTGGGATAAAGGCATGCTCCACCATACCTAAT
CCAACCTCACAATTTTTTAAGTGTGTATTTATATGTGTGTGTGGTATATG
TAAAGGTGTGTGTGTTCATGCACACATGTGCAGAGATCAGAGGAGTCAGG
TTTTCTCATCTATCACTCTCTGCCTTATTATTTTGAGACAGGGTCTCTTG
TTCGATATTACATATACTAGGTGAGATAGCCCAGGAGCTTGTAGGAATTC
TCTCCCATTTCTACCTTCCAAATGTGTGCTACTGCATCTGGCTTTAAGCA
AGTTCTGGGAATCTGAGGTCAGGTCCTTACACCTATGTAGCAACTCTGCC
TACTGAGTCATCTTACTAGTATTCACAAGGTCAAAGGTTGGGACCAACAG

FIG. 3D (26)

CCAAGGTTGTCCTCAGATCTCCACACAGATGTACCCACAATTATACAAAC
ACTCAACATAAACCTATTTACACACCCACATCACACGCACACACATACAT
GCACATACAAAAAAATGCTTTTTGAAAGAAGTAGAGAATGCTAGATATGG
TATTACACGTATATAATCCAAGCCACTCTGGAAGCTGAGGCAGGAGGATT
TCAAGTTTGAAACCAGCTTGACCACATAATTATACCATGCCTCAAAAATT
GTATAGAGAATAAGAATGAATATGAATGAGACTAAAGTCATATCTCAGTT
ACTTTTCTATTGCTGTGGCAAAACACCATGACAAAGGTAATTTACAGAAG
AGATTATTGGGGCATATAGTTTCAGAGGGTGAGTCCATGACAATTATGAT
ATGGCACTGAAGTAATAGCTGAGAGCTTAAATCTGGTCCACAACATTAGG
CAGACAGAGAGCTAACTGGAAATAGCCATGAGATTTTGAAACCTCAAGCC
CCACTCCTAGTGATGTCCCACACCTCCTAATCCTTCCCAAACAGTTCCAT
CAGCTGGGAACAAGATATTCAACATATAAGCCTATGGGGGTCATTCTCAT
TCAAACCACCAGTAGTAATTATTAGAGCCCAGCAAAGAAGGAAGGGATAG
AAAGAAATGATTGATGGGAACTGGGGTGAAGTCTGATACAGAGAGATCTT
TATGTACTGCAGCGTAGCTCAGGAAGATAACTATGGTTAAGGACAATTAG
CTAAGTGATTAGTAGAGAGGATTTTAATATTTCCAATACAAAGAAATGCT
GCAGGCCTGAAATAGGGTACGTTTCAGTGACCCAGATCTGATTATTACAA
CTCATACACTTGTACCAACCACATAAATATGTACAATAATTGTGTCAGTT
TTATATTAAATAAAAATGTGGAGCAAGTTAAAAAATGCCTGTTTTAAACT
GATCACAGTTATATGCCAGCTTTTCTTTGCTGTGACAAAATACCATAGGG
AGTAGTTTATAAGGAAAGAGATTTCCTCCAGCTCATAATTCCAGAATTTT
CAGTCTAGAGTCAGTTAGTTCTATCATATTGGGCCCACAGCTAGACCAAA
TACAATGATGGGGAGAATGTGGTAAAGAAAAGTATTTACCTCAGAGTGGT
CAGGAGGAACACAAGACAAAATATACATTTCAGTCCCATACCTCCAGTGA
CTTGCTTCATCCAAACAGACGCCACCATCCAATAGCCATTAAAATACAAG
TCAACCAGTTGATTGACATCCATTGATCTTAGTCATATCCCTAAATTCAA
CCTCTAAGCTCTGATGCTCTGGGGGCCAAGCCTCTATTGCATAAATCTCT
GGAGCATATTTCATAATATGAAATATTAAACAGGTCTCTCAGGAGCTGTT
TGGTAGACTTAGTTGTTTTTTTTTTTTTTTTTGTTTAAGGTTTTTTTGGTT
GGGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTGTTTGTTTGTTTTTT
TTCGAGACCGGGTTTCTCTGTATAGCCCTGGCGGTCCTGGAACTCACTTG
TAGACCAGGCTGGCCTCGGACTCAGAAATTTACCTGCCTCCTCCTCCCAA
GTGCTGGGATTAAAGGTGTGCGACACCACTGCCTGGCCTAGACTTATTTT
TTTAATCAGATTTGAGTCTTTGCCTCTGGAATCACAGTAGCTTTTCCCAT
TCAACACCTAGTTTACAGAAGAAAGAAAACCCAATTTTTTTTTTTATAAT
CATTAGACAACTAGAAGTTTTCCCTCCTATTAAGAAAACATATTAACGGG
CTGGCGAGATGGCTCAGTGGGTAAGAGCACCCGACTGCTCTTCCCAAGGT
CCAGAGTTCAAATCCCAGCAACCACATGGTGGCTCACAACCATCCGTAAC
GAGATCTGACTCCCTCTTCTGGAGTGTCTGAAGACAGCTACAGTGTACTT
ACATATAATCAATAAATAAATCTTTTTAAAAAAAAAAAAAAAGAAAAAGA
AAAAGAAAACATATTAACAGTATTGAGAAAACTGTTGGCTTAAATTTGAT
GATTTGAATTTTATTTTACTAATAAATGCATGTATTGCTGGGCATGGCAG
CACATCCCAGCACTCAGGATTCCGAGATAAGAGATCATAAGTCCACGCTA
GCTGGAATAGCAAAATAAAATCTTTTTTTAAAAAATATACATACATACAT
ACATACATACATACATACATACATACACACACACACACACTTTTCTCAGT
AGTACGGCCAATTAGTTGACTTGTCTAACGGAGGGAGGAAGAGGAGGCAG
AGAGCATGCTGTTCAGATCACGTTCTCTTTTGCATTCAGTCTGGGACCCC
AGCCCATAGTGTGGTGCTGCCCACATTGATTATTGGTATTCAGTTAACCC
AGTGTAGAAACTCTCTCAGAGACATGCCCAGATGCTTGCCTCATAACCAC
TGTGTATGTATATATGCTTACAGAAAATATACTCATCATTACACATAAAT

FIG. 3D (27)

TTCATCCACTTACCTCTTATGAAAAGTTGATTATTTACTGAGATTTTTCT
CATTCTGAAAATCCATAAAGTCTACCACATTGATTAAATTACTTGTTTTT
TACCTGTTATTGCTCATGTTAGAATTGCTTTCCTTATTTGGGGTAAGCTG
TCGTTGGCCACTGTGAGGGGCTTATCAAGAAGTCAGAAATGGGAACACCT
TCTAGGAAGTCAGGACTGGAAGCTTAGCTGAGCCAGCAAGTGTTTCTCAC
ACTGCACTTCCTGTGAGCCTACCTGTGCGGCATCAGGAACTGGAGTTGGG
ACCTTGAGGATTGTTCCCTGGAGGCAGGGGTGGAGTCAGGCAGGGGTGAA
GCTGACTCACAAGATGGTCTTGCCTTTCAGTTGTTTCCTCTCGCTGCTTC
TGGTGGCTGCAGTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAGG
CGGAGAGAGGTAAGCCCAAGTAGACAAACTCCACATAAAACTCATTTTTT
TCCTTCTTTCTAGGCAGATCACTTTTACCTGTTGAGTGATGACTAATATT
CATATGAGAAGCATGCTGTTTAACCTGCATTCTGTGGTTCCACTATGTGC
CATCAGTAGATTTTAATTATTCTTGCATAAAGTGTCATTAGTTTTGCCAC
TGCTTGATTCAAGTCTTCCTAAGAGTCTTTCCTAAGAATATGAGTGTAGA
GACAAGTTCAGCTCAGTGACAGAGCACTTGCCTGGCATAAACTGAGTCCC
TGGATTCTAGTCTCAGCACCCTCTAATAGCACAACACTAGAGACAAAGCT
TCTAACCTGTGGGTCTTGGGCAGCAGGTAGGGGAGGGGGATTTAAAAAAC
AAAAACAAACCTCTAGCTGTAGCCTGTGTCATTTGTTATGACTAAGCACT
AGAGTGGGTACTAGTAGACATGCCATGTGGACATTGAGCACCTCTCCATC
CCAGGCACTGATCCAGGTGGTTCTGCTTTATCTTCATCTCCACCCTAGGA
TATAAGGGAGGCTACGTAACTACCCATCACCACACAGATGCTGAGGTACA
GAACTGAGGGGTAACTAGTGCCTCTGCCTTCACAGCACAGGTTCCTAAAC
ACGTTTTCTACAAACACTTCATTTGTTCTAGTCTGTTCATTTAAGAATCT
CATGTTCTGACTGAATGAGCTAGACAACTCACCCTAGACTATACATTCTA
AAGAAGGGCAACAAGGCAGTTTTGTTACTGTTGAGAAGAAAACAAAGTTA
TTTCCGTATGAGTTATTGAGATAGAATAGTAGAGATTTGTCTGAATACAA
AATAGAAAGTATATAAAAGTATATAAGTGGATCATAAAGAAAGCAACAAT
CAACTGGAAAATATTTGCAGTATCATGAGAGAGAGAAAACTAGAAGATGA
ACCCCCTCAAAAAAGGATTTTTAAAATATGCTTAGACTGTATTCAGTCAG
CTAATAAACTTTTTTTACCTTTATTTGGAATTTACGAATAGCACTGAACC
TGACCATTGTAAATGCACGAGGTCAGGCATGACTTGTTCCCAGTAGGAAG
TTGTTTTTAGTTCTTGCTGTGGCCTGGGTCCTGATGGAAGTTCTTTACCC
ACCTTATCTCCTGTCCTCTTGGCAGAGGTTCTAGAATAGTGCTGTGATGG
GGTAGCAACTGTCTTCCTGTGACCCTGCACCTAGATTATTACAGAACCCA
GACTGGGTTTGCTGAGTTAATGGAAATTCTTTCTAGGTTCAGTAGAGAGA
TGTGCTGACACATACTAGGCCATCTAGTTTTTCAGTAATGCTCAGAGACC
GCAATAGGATATGTAACAGCAACAAAATTTTTAACATAAAATTTCCCTTC
TAAAACAGAGTGATGATTTATGTAGCTTCAGGATCCTGCCTCCTAGAAGA
TGGTTTGAAGCAAGGCCAGTTTGTCTTCCCTAGCATAACCTCAGAAGACC
TCTCATATTATTGATGGTATAGGAATGAATGCCCACATTCTGTATTTGAG
ATGTGTGCTATAGTATCTCATCTGACCCAACATGAAAACATTTCAAGCCA
TGTGTGCTTGGGTAAGGTAGGAGTTCAAAGTCATCCAGTGAGTTCAAGGC
CAGCCTGGGCTGCATGAGACACTGTCTCATAAACAGACACTTGAATCTCA
TTTAAAGAAGACATTGAAGACTTGATACTTTGAACACCTATCCTAACGTA
TCCACCCCCAAATCCAGAGTCCTTCATGTTCTTGTCCTCTGCAGTTCCAC
TTTCATTGTGTTCTCAGCAGCAGCTCTCTCCGAGGAGAGTTGTCTCCCAT
CCTATCAGCCATCTTTTTTATTGTTGTTGCTCTGACAATGTCTGGTTCAG
GTTTTAACACAAAGCAAGCTAGAGTGATTTTAATCTAGCAACAAAAATAT
AAAAAGGTAAGTTTTTGCCCTTTTATATATTCAATCAACAGATATCATAG
CATTATATCCTCCACTTTAACTTTTATTTCTTACTGGTAAGGGCTTTTTA

FIG. 3D(28)

TAAAAATATAATAGTGTTACCACATGTAACAAAATTTGATACCTTGTGCT
ACCTAGCACCTTGTCATGTCCAGTTTTCCTCAGCTGTCACAGAAGCGACA
CTGCATCTGATCAGTTTGAATCAGAGAGAGTGTAGCATGTCTAATATCTA
GTATTCACTAATAAAATCTCAGTACTAAGCATATTAATAATACTATATTA
TTCATTAGCAACTTCTTCGGGAGATGCAACAGATGGCCAGCCGCCCCTTT
GCTTCTGTAAACGTTGCCTTGGAAACAGATGAAGAACCTCCTGATCTCAT
TGGGGGAAGTATAAAGGTGAGAAGTGGCTCAAAGGTCCATATAGCTTTTC
AGAACTCAGGCCTCAGTTTGCTAGGCTACAGACAGCAAGCGCTCTGTGTG
TCACTCCTGTCTCCTCTCTAACAGTTAGTCAGCAGAAGCAACCCCGAGCG
ACCGTAAGGGGCTCTGTGTGTGGCTTTACTTTTCGAGTTGTTGCATGTCA
GATTTTAACATGCAAATTAAGCTTGTTATTCTTACTTTGTGGCATAATAC
TTTATAGTTTTTTATTTGGAAATATCTAATCTGGGCTAGGTGTGATGGTGC
ACATCTTTAATCTCAGTTCAGAGGAGGCAGAGACAGAGGCAGGCAGGATC
TCCTTGAGTTCTAGAACAGCTGGTCTACATAATGAGACCCTATATGTTAG
AAAAAAAGAAAGAGGGGGTGGGGGAAGGCAGCTAACTTTAACCATTAATT
GAACCAACACACACACATTTTGTTCAGAGCCCCAGTACTCAATTAAAAGC
CAGGCAGGCATGGTAACAGTACTTAGGGAGTCAGAAACAGGATTCCCAGA
GTAAGCAGTCTGACTAGGCTAGCAGGAAATGGTGAGTTTCAGGTTCAGCA
AGAGGCCCTGCCTCAGTAAGTAAATTGAAGAACAACTGAGGGAGACTTGC
ATGTGCACTTGTGCATGCACCCACACATGCACTTGCACACATACCATATG
TCACCATGCTTAGACTATAAAATGTAGTCACTACTGGCAGCACATGCCTA
CAATACAGATGCAGGAGAATCACTGCAAATTTGAGATCAGCCTGGGCTAC
TGGACAAGATTTTGTCTCAAGAAAACTAAAACAATACAAAAGTGTACTGG
GGGGGTTATTCTAATGCCAGTGTTTATGACAGCACATTCAGAACTGACAG
TAAAGGCAATCAAGGACTGTCAGTGGTGGGTATATACATAGGCAGAGGAG
CAACTGCTACTAGAAACTGTTTATCCTTTAAAAGACTAATGTATGCTGCA
GCATAGACAAACGTTAAGTTGTGTTAAGTAAAAGATGCTGTATCATTCCA
CTTACCCATCGAGAATAATCAAATACAAGACAGAGTAAAATAGTGACTGC
TAGAGGCTTAAAAGAAAAGACCAGGGGGTGGGGAAAGGGAGGGAAGGAAG
TGGGAGAGGGAGGAAGGGAGAGAGGGAGGGAGGGAGCCAGACTTTGTGGC
TTACAGCATCAAGAGGCTGAGGCAGAAGGGTTACAAATTCAAGGCCCTAC
TGGGCTACATAGTGAGAAGTAGGATTTCCTTGAGCTGTCTTTCTAGGTCA
TAATCTCTCATTGGGGGAAGTCAGGGCAGGGACTTGAGGCAGAAACCATG
GGGAATGCTATTTGCTGGCTCCTTCCCAGGCTCCTCTCTAGCTTTGTTTT
CTCATTTTGTTTTTACTGTCTATGGGTGTTTTACCTGCTTGTTTTTCTGT
GTACCATATACATGCCTGCTACCCACAGAGGCACTGATGCCTGGAACTGG
AGTTACAGATGGTTGCAGGCTGCCCTGTGAGTGCTGGGAACTAAACTCGG
GTCCTCTACATGAGCAAGTGTTCTTAACCATTGAGCCATCTCTCCAGCCT
ATAAAATTCTTTTTTAAAAATAAAGTCTGCAACAGAAAATGAATATTTTC
TAGAGCTGAAGCATTCAATGAGTGGATAAAGAATCCATTTGATGAGCTAT
CTACCTTTCACAAGCTCTTAACCCCTACAGACTCAGGACTTAGTGGCTGG
AAGATGAATGTAAAACAGGTAGCTCTCTCCATAATATCTGGTCTGTTTGT
GCCAGGTGTGCAGAACTGTGCAACAGGTCACCATACAAACCGGCGTGGGC
CTTTCCTGACACTCACACAGCTCTCGGGACAGTGCCCGTGGGGACCTCTT
ATTGACCTTATAAGCACCTGACTGTGCAGTGTAGCAGGGAGTTAAGGTGC
TTCTGTTTTCTTCCTCCAGACCGTTCCTAAGCCCATTGCCCTGGAGCCCT
GCTTTGGTAACAAAGCCGCAGTCCTCTCTGTATTCGTGAGGCTCCCTCGA
GGACTGGGAGGAATCCCTCCTCCTGGTCAGTCAGGTGAGTAGACAGGAGA
CAATGACAGATATTGGTCTGTGAAGGACTGAGTCTTAGACACTTCTTCTG
GTATAGAACCTGGGTCTGGGCACAGTGCTTAGTGGTACAGAGCTTTGGTG

FIG. 3D (39)

GAACAATTCTATAGTCCCCAAACTGTGTTCTGAGCACTGACATTCCTGTC
CTGGGGTGGAAGTTCAGGACCTTCCTCACGGTGCACAGCGTCCTCAGACA
TTCATGCTCTGGTCCCCTTGACTCTATTGATCCCTGCTTTCTTTTTTTTT
TTAACCCCTTGTTCTTATCTCAAATTTAGGCTTTTTCTTCCTTGATACAA
GCTCCTATTCATCTCCATGCCTCTGGCTTCCAGCCATGTCCTCAAAGCTT
GTGTTGCCAAGTACAGAGTTCTAGTCATGCTCCACATCTTCTTAAGGTCT
TGCTATGCAGCCTTAGCTGGACGAGTGCTCGTTATAGGCCAGGCAGTGGT
GGTACACGCCTTATGTCCTAGCACTGAGGAGGCAGAGGCAGGCAGATCTC
TGAGTTCAAGACCACCCTGGTCTACAGAGTAAATTCCAGGACAACCAGAG
CTACATAGGGAAACCCTGTCTCAAAAAAATAAAAACAACAACAGGAACAA
CCCCAAAAACTCATTATATTGCCCAGGCTGGCTTCAAACTCATAGTTATC
CTCCTACTTCAGCCTCCAAAGTGCTGGGATTATGGGTGTGACCCTTCATG
CCCAGATTGTCTTAAATATGAGGCATGAAGAAGTATTATGAAAACATAAA
GGATATTTTGAAAATTATAATTCTACTGGGTTAATGCAGATCCATTTTCA
TTTCATTGAAATAATGATACAGCCTTTGGAGGTTAGGGGAGCCTCTCCTG
TTTTCAAACTGACTTTGAACTTCTGATCATCCCGCCACCACCGCCACCTC
CTCCTCCTCCTCCTCCTCCCCAGTGCTGAGATACATCACTACTCCTGGTT
TATGTGGCACAGAGGCTCAAACCCAGGGCCTCATGCATGCTAGGCAGACA
CTCTACCAGCCAACCTACCCACAGCTCCTAGATGTGCACCGTATTACAAA
CATTTATTCTTCAGCATGTTTTTTTTTTTTTTTCCTAAAAATCATCTCTA
CAGGAAACAAGTACCAGTGGTGTTTTAGGGCAGGAATAGGAAGAAAATAT
TTTTACTATATACTCTTTTTTTTTAATCATTTTTTAGATTTTATTTATTT
TAAAATTTATTTACTATTATTAATAAGTACACTGTAGCTGTCTTCAGACA
ACCCAGCAGAGGGCATCAGATCTCATTACGGATGGTTGTGAGCCACCACG
TAGTTGCTGGGATTTGAACTCAGGACCTTTGGAAGAGCAGTCAGTGCTCT
TAACTGGTGAGCCATCTCTCCAGCCCCTACTATATACTCTTTTAAATGAC
TTATTTGCTTTTATTTTTATGTGCATTGGTAATCTGCCTGCATGTATGTC
TCTGAGAGAGGATCAGATTCCTTGGAATTTGAGTTACCTTGTGGGTGCTG
GGAATTGAACCCAGGTCCTCTGGAAGAACAGCCAGTGCTCATAACTGCTG
AGCCGTCTCTGCAGCCCCTACTATATACTTTTTTTTATAGTTTTGAATTTT
TTTTTCTTTTTTGGGTATTGCTAAGGATCAAATATAGATCTACTATTTATT
TTTTATAACATCCATTAGTATTTTTATAACTTACTACATAGTTTGCCAAT
TCTTTTATACATGTCCATCAAACATGTAAGTCATAATTTATATAAACCTT
GTGTTAAAGCTGGAGGCACAGAAGGAAGATTGCTACAGAGTGAAGTCTAG
ACTAGCCAGGGCTATATAGTGGGACCCTGTTGCAAAGAAAAAGTTCTCTC
TTTAAACACAAAGGCAGTATGAAAAGACATACCTTGATTCTGAAGCTGTG
CATAGGAATGCCTCACACAGTGTTCTGCTCAGGACTATACTCAGATGCAG
TGGTCTGAGGGACTTGGTGGTGTCTCAGCCAAAATAACCTGGAGTTTAGT
AGGAAAGTCTCCTTTATCCGTGTCCAGTCCTGAAGGGAAGCCTTATTTAT
GTATGATGAGTCAGGACCCATTGTCTTCATCTTACTTGGCATCCCCCCAG
CACTGAGTCTCTGAGTTAGCCTTACTTGGACAGAGTGACTCTCTGGGCAC
TCTGGACAGCATCTCCTGCTTCAAAAGGGCAAGATCTTTAGAAGACACAG
AGATGGAGCAGGTCTTACATGGAGATATAGCAGCTTTTCCTTCCTGACCC
TTGACCCAATGCTTCTTTGGAAATCCTCATGAAACCCTGCTCCTTTCTGG
AGACCCACCCCACAGCAGGGTTATCCATGCCAAGCTTCCTGTACTTTCTC
TTTTTGAGGAAGCACATACACACAAAGTTTTAGTAGCTCGCACATCTCAC
TGTGAAGTAGTGATACTTTCATTGCTATCTTCTGGAAACAGGCAGGAGTA
GGCACACGCTCAGAGCATAGCTGCACTCTCATTCACTTGCCACCCTGAGG
CAGAGCACACGACTTTGTGATCTGCTATGGAGGAGAGAGAAATGAGTAGT
TAGGTGTGTATAAATAAGCTAACACCATCACCCCTTTATCTTTCACTAGG

FIG. 3D (30)

GAAATGTAAAAAGAAATCTGAAATTATTTTGTAAAAAAGTAAGCTGCTTC
ATGACACATGTCCCCTCTTGTGGGTTCTTCCAAGGTCTCGCTGTGGCCAG
TGCCCTGGTGGACATTTCTCAGCAGATGCCAATAGTGTACAAGGAGAAGT
CAGGAGCTGTAAGAAACCGGAAGCAGCAGCCGCCTGCACAGCCTGGAACC
TGCATTTGATACTGGGGCAGGAATTCGCCCTCACAGAGGGCGTGTGGTCC
ACGAAGCTGTCTACAGGGGAGGCTGCAGGCAGGAAGCAGGCGTGGGGCAG
AAGACTGGGGACCCTTGAAGCGTCCAACTCATGTGCATGATCATGCAAGC
TGTTTTCATGGCTCACCCCTCTGTGTCCAGCATCTAACCTTTTACTTCTG
TGTAGGAAATAATTTAATTACAAGTCCAGGAATGGTCTGCTCTACTCATG
GGTGGAGGAGACCAGTGCCGACCCCGTGAGAGCTGAAGGTGATGCTGAGG
TCCCTTGTGGAAGCCTCTCTTGGGAATCTCAACTGCAGAGGAGCTGCCCT
CTGTCAGCAGCTCTCCAGCATGGTCCTCTGACACTCCTCAGATGAACTGT
TCTCATCGGAAGCTTGCTGTCTTTTTACAAGATGAGCTTTTACTCTCTTC
CAGGAAGTAGCTTTTTTTCTAGCTGAGAATTAATAATGGTCTTTCTCTTT
GGAAGTCATATCAAAGTATAATTGATGGGGGCCTTGTTTTGTTTTGTTTT
GGTTTTTGGAGACAGGGTCTCACTGTGTAGTCCTAGCTGGCCTGGAACTC
ACTATGTAGATCAGGCTGGACTGAACTCACAAAGATCCACCTGCCTCTGC
CTCACAAATGCTGGGATAAAAAGCATGAACCACCAGGCCCAGCAAAGAGG
GCTATTCTAAATGTCAAGGTCAATGGAGTTAGAATATATATAAAAAAATG
CAATTGATAATTCTCTATAGAAACTTGATTAATTTTAATCCATTCTTTCC
TTCTCTTTCTCTCACTCTGTCTTACACACATGCACACATACACACACACT
AAGTGCCTAGACTTTGAATAGATCTAGCAATTGGACATTAGTAAGCCTAA
GTTTTTACATGATTGCATTCCTACATTCTTGTAAACTTTAAGTAACTACC
ATTGCAGTTTGTTCTTTTTTTAAAGTCTAATTTGCAGCCAAGAACGAGTA
ATTCTCACCCCAAGCAACATCTAATAGGGACTGAGTGACCCCAGCCCAGC
CTAGTGTCACTTTAGGCCTGACGTTTGAGCAACCCTCGGCTCTTGCCAAG
GCACCACAGAATGCACTTGCTCATGCCCTGTGCCTCTTGAGCAGAAAAGA
GCACTGACAACTGGGACACCTGGCTCTGTCTTCCTACAGCTGCTCGCACT
GACCTGTGGGAACCTGTGGGTCATCCCCAGGCTGAATGGAGTACACACTA
GAAGAGGGATGATGCCTAGCATTGGGGCAGCATCTGCTCAGCACATGGAA
AGGGACCTGGTTCCATCTCCCCTGGGCAGGAGTTGGTCCAGCCTCCTCCC
AGACCCAGCTGGTGGCTGTGAGGAGGTGGGGAATGCTAATGAGAATGAAA
AGCACATGGGTTGATGGGAAGGGACAAGATTACCACGTTAGGAGGGTGAG
CAGCCCTCTGCTATGTGCCCAGGACCCTGCCTGGACATTGCATTTCCCCA
TTTATGGTGCTCCGTATTCTGGCATTATGCAGCAGCCTCACACACCTGTC
CTCTCCTTCTTCATGTCCTACAGTTCTGCTATCACCTGACTAGAATAGCC
CTCTAGGCAACAGTGCTCAAATGTATGAGTTTGGAGAAGTTAACAATCAG
AAGAACAAAAACTGTAGTGTTTCACCTTTAAATGCAGTGTTGAAGAGGGA
GCCTTTCTCTAAGCCCTGCACTAACCCACTCCTCCCAAGACTCTTGTGGA
GTGACAGTTCCAAGCTGAACCATAAATCACTGATGCACAAAACACTGCTA
GAAGGCTCACCTCTCAAAACACGACTCTTTGCATCACTATTAAAGAGCAG
AAAGTTCTAGAAATGATCCCAGCCTCATCCCCTATACAGTTAGGAGCTCC
CCACATCTCTACCAAAACCCAGCACATAAGTATCTGCGTGGTCTAGCCTT
TCATCTCCGTAACAAGCCAGGGGACTCTTGGCCAAAAGAAAGAAAGGGAA
GTTGCACTAGGGCTTGTCCGTCCATAAGGAATTCCCCTCTGCTTTGCTCA
AAGGACCAAATTTCTTTGGCCAAAGAAGTTGCTTCTATGTTAGTCCCATA
CCCTGAAGTAATATGTACCATGGCTCCCACCTACCTGTTTATGCTCTCCC
TGCCCCCAGGGAAACTGTTTATTCTTTCAAAAGAAGCAAACAGCGTTCAT
TTCTGCTCCTGTAATGGAGAAACAGCCAGCTCCCCTGCATCCCTTACAGC
CAACAGCTCCCTTCAGGCTTAGAGCAGGGGGAATGGCAGGGATTAAGAGC

FIG. 3D(31)

TCAGCTCAGAGCCAGTTACCAAGATGGAATGGAGTTGTGACCCAGTAACT
GTGTCACGAGAGACCATGTATATAAAATAGTCATGACGACACTGACCTCT
TGCACTTGTACATAACTATACTGTAGTGTCCAGAATGTTCAGACATTCAG
GGTGTACATAAACAGAAGAGTATCATAATGTATTTTTATTAAACACTAAC
ATCTGAGTTTCACCTAATCTGTTTCTGTGCCATATACTGGGTATCCAAGC
TCTGGGAAGTTATCCTACCAGGCCCTGATCTGTTGATAAGGCACTATACA
CCATGCTGGTGTGTTCTGTAGCCTTGTGCCCATTAGGTAACTGAACAATG
ATTCAGCTCTTAGAATACCTAGGAAGACAGCAAGCAGGGTGACACACGGC
TGTGATCTAAGCATTCAGAAGACAGAGGCAGGAAGAAAATTCAAAAATGG
GGCTGGAGAGATGGCTCAGTGGTTAAAAGCACTGGCTGCTCTTGGTCAGG
ACACTAGTTCAGTTCCCAGTACCCACATGGTGGCTCACAACCTTCTGTGA
CTACAGTTCCAGATAACCTGACACCCTCCTCTGGCCTCCTCGGGTGCCTG
TGGTGGTCCACCTGGTGCACAGACAAACACCCAATACACACAAAACAAAA
GTAACTCAAGAATAGCCTGGGCTACATAGCAAGAGCCTGTCTCAAAACAA
ACGAACCTATGAAGAGCCAGGCAGTCTATCTATTTACATGGCAGTATACT
AGAGAAACTCAGGAAGCAAGAGTGTTCATCACTGTTGTAATTTCAAATGC
TCCTTGTGATTTCTGGCATCTCTGTGGGGTGAGGTGTTCTGTTACTCTTC
ACATTCAAAGACTGTCACCCATGAACGTCAGACTTTGCAAAGGGGCTCTC
TAAGCTGCACTGTTGTGGCTTTGTCTAAAATTTTAATGACGTTTCTGAGA
ACCATGTTCTTTTTATACTAAAATCTGGGGATGGGAGGGCTCATTTGTTG
ATAAATAGCACTATTTTCCCACACCTCAGCCTCCTGTCCCCGTCCTGGTC
TTCCCTACACAGTCTGGAGAGGGCTCTGAAAGGTCCACAGAGTTTGACAG
ACACGAAAGCAACCCATTGCCCCGTTGACCTGACCTGGAAGAAGACTGTC
AGCAAAAGGAAAATACCAGAATATCTGGAAAGCTTGAAGTGTAAGATGGG
ATCTCGTTGGGGAATTGGATGAAGAAAAGCAGAGCGCCTCTGGTAGGTGA
CTCTGCAGCCTGCCAGCGCCCGCCCTCTTTCTACACAGCAGAGTGTGCAT
GGCAAGGAAATGAGTCACCTCCTTGGGGGATGGTGCTGTTTTTATGAAAA
CCTCTGATCCTTGGTGTCCTTTAATTGATCTGTTCAACAAATATTTACTA
AACACTTCTAAGCTAACATTAGGGCAGTGACTGAGGTGGAAACCCAGCTC
TTTAGACAGCTGTCATCCTAGGATAGCTTCCTGGAAGCAGAACCAAGAAG
CCAGAAGGTTCTTCCTAGGGTGGCCTTGGCTCCCTGAAGGAATCTGAAAT
GCTGACCCTGTCACAACCTCCCAGCACAGCTTTGGAATGAGACATCAGCC
TGGCCTCCAGCAGAGCAGAGGCTCTGGAGCTCCACATCCTGCCTGCAGGG
AGCCCTCAGGGTGCCCTCCAGAGTACAGGGAGAAACTAAAGGCAATAACA
GAAGCTGCTCTCAGAGCCTGACTGTGCACAAAACACTAGTGAAGCCTGCT
GAACTAATTCTGCCTCTGGAAATCTTTTCTGGTTCTTTACAGTTTGTTGT
TTTGTTTTGATCCAAGCTTAGTTTGTTACTATGTGTGATTTAGCATCTGT
CGCACTTGTGTAAATATGGAGTAAGTATTGTAAACTATTTAATTGCTGCG
ATTGTTGGGTTATACATACATTTAGGACTGCAATTTTTTGGTATTTTTTG
TATTGTAAAATAACAGCTAATTTCATCAGGAACAAGAGAATTAAGGGGGT
CTGCATTTTAAATGCAGATGTGAAGCACTTGTATATAAATAAAAGTAAAT
ACTATAATACAAAGTTCCTTCTGAAATAAAAGTAGATCTGGTAAAAATGT
GCGTGCGTTTCGTTCTGAATGTTCAATGCTAATTTTGTTTTATTTTATAT
TTACATTTTAGTCCTTATTTTAGCAGTGAGGAGACAGGCACAGCAGTGCA
TTCTCACCTTGGCAGCTGAGGAATCCCCTAGAGTAGACTGCAACTCAAGA
CTCTTGGCTTCCACACTGAAAAGAGTTTCAGTTTATGAAGCAGAGTTTAG
GAAGTTTAGTGAGGAATTTAAGGACTTCTTTTAATGTTTGTGTCTACATA
TGTGGGTACATATATGACACAGCATGCATGTGGAAGGCAAACAACACCTT
AATGGAAGTGGCCTGAAGAACAAACTCAGGACTTCAATCTTGGCAGCATA
AACCTTTACCTAATGAGTCATCTCCAGTCTATACGGGGTGTGTGTGTGAA

FIG. 3D(32)

CACATGTGCAACAGCACACAGTGGAGGTCAGCACAACTCTTGCGAGTCAA
TTCTCCCTTACCTTGTAAGACCTAGAATTCCACATTGCCCAGGCTCTGAA
AGTTAGGTTGGGTCCACACTGGGCCATGGCTGATGAAATGTTGGAAAAGT
GATAACACCAAACTTTTGCACAGAAAATATTTTCATCTGGGGCCTTCCCT
GGAGTTCACAGGCTAAAGTGTTGGAAGGAACATGGGTCCCTGAGCCACCA
CTTTCACAAAAACTACCTGATCAAGAAGAACTATTCTGGGTTTCTGTTGC
TAAAATTCCTTCCCAGAGAGAAATGTAAGCAATGTCTGCCCCTTCAAGGG
TCCCAGCAAGAAACCAAGGCACAATTCCACCAAAGTTCACTAGAAAACCA
GTGAGTTTATTGGGCTTCCGTGCAGAACATACATGAGGGGTTACTTAGAG
AAGTGTGGATACTCCTCCCCCTAACAATCCACACCCTGAAAAAGCCTTAC
CCAGCAGGGATGAGGGCTTCCCCAGACCCACATTGATGGTGCTCCCATTC
CATTTTTCCCTGGCATGCAAAGAGATAGACAGAAAAATAGATTATATATA
ATATACACATAAATTAGAAAAATAGATTATATAATATACACATAAATTAT
ATATTATATATATAATATATAATACACAGATAGATTATATATGATATATA
AAACACACAGAAATAGGGTATATATAATATATAATACACAAACTACTCAG
CTATTAAAAACAGTGGATTCATGAAATTCTTAGGCAAATGGATGGAACTA
GAAAATATTCTGAGTGAGGTAACCCAATCACAAAAGAACACACATGGTAT
GCACTCACTGATAAGTGGATATTAGCCCAGAAGCTTGGAATACCCAAGAT
ACCATTCACAGACCACATGAAGCTCAAGAAAGGAAGATCACACGTGTGGGT
GCTTCTGTTCTTCTTAGAGGAACACCCTCATAAAGTAGTGGTGGGGGGTG
GGGGGAGACAGAATAGGTGGTTTCCAGGAGAGGAGGAAAACAGGAAAGGG
AATAAATAACATTTGAAATGTAAATAAAGAAAATACCCAATAATAAAAGA
AAAAGAATTTTGAAACAGAGGGTAAAAAATAATACACAAACCAGGTAGAT
AGATTATATATAATATATATAACACAGAGATAGATAGATAGATAGATAGA
TAGATAGATAGATAGATAGATAGATAGATAGATAGGTCAACTGCTC
GCCCCTCCACTAGGTAACATGCAGTTAAGGCAGAGCTGCATACAACAGAT
GTTAGGGATACTCAGGTGAGAATCTCAGGCTTTGCTCCATCCATCTATGC
TGGGGTGTAAGCTGTCAACAAGTTTAGCTGGGATGATGCTTTGCAAGAGG
GCACAGCTGAATGCCCTAAGATGGTAGATGCTTGGCTCAAAGGAGACACT
ACAGCTCTGCATCAAGGCAAACTAACTGAGATGAGGGCCTTTATTTTCCA
GATCTGTATCCTGGAGCATCATTCACCTGTTACTACACTGAAAACATTTG
GTGTTGGTTTCATGGCAGATGACAGGCAGTGAGAGAAGTACAGCAGCGGA
CTGCTAGAGGTGGGGGTTCTGTCAGGACGTGGGAGGCTGTTTGGTTAGTA
ACTTGGAAAGCAACAAGTTTTTAGCTAGAGGGAGAAAAGCTGGAGATAAC
TGTACTTGCTTGATTTCTTAAATATCAAATTTTATTTTATGCATATGGGT
ATTTTGCTTGCATGTATGGCTATACACTACATGCTTGTGGTGCCCACAGA
GACCAGAGGAAGTAGTGTGAGCCTCTGAAACTGAAGTTACAGACATTACG
ACTTGAGTGCCTGAAACTGAACCTTGGTCCTCTGGAAGAACAGCCAGGGC
TCCTAACCACTGAGCTATCTCTCCAGCCCTGACAGAACATCATGTACTCC
AGGCTGGTCTCAAATTTGCTTTATAGCCAAGAACGGTCTTAAATTCTGAT
CCTCCTGTTCTCTCAAGTAGTGGGGTTACAGGTCTACACTGCCGTTTTCT
TGAGCAAATCATTACAAATTGAGTTCTAAGCCAGGTGTAATAGTTCATGT
AGTAACAATCTGGAATTTTGGTCTCTTAAAAAAACAAATATTATAAGAAT
GTATTTTCATTTTAATCCCAGGTGTATGGCATATATCGAACTGCTTTGGA
CTGACTACAGCAGCTATGATTTTTTCTTGTTCTAGCAGAGGTATGGTTTT
GCCAGCTACAGATAGTTTCTGTGATTGTGTGACATTTGGAATTCTGGAAA
CTTTTCAGATGGTATATAAATATTAGAGCCCCAATAGGCAGAGTTGATGA
TTGTTGGTCATTCAGGGGTATTGGTTGTGGTTAGTAGTCTTGCTTGAAGA
AGAAACAAGAACAAATTAGATTCAGAGATCTCTATATCTCTCTCTATCTT
CCTTTCTGTCCTATCTAGTAATAGGGGGTAAAACCAGGATGATAAAGGGT

FIG. 3D (33)

TGGGGGAACCCACAAAGTAACAAAGACTGGCTACAAGTGGCACCCAACTT
GGAACTCAAAATTGCCATAGAGGAAGCAGCAGGGGATGAAGGAATGGATT
GTGGCTGTTGTTGCTGGGATATTCCTCACTTTGCTCCCAGAGGGGATTTT
TCTGAGGTTTTGTTGTTTTGCTTTGCTTTGCTTTGGTTTTCTTCTACATA
TTCTGTTTTTTTAAGTAAGTTGAAATAATAGCCGAGAAGCTGGAAAAGTTT
GGTGTGGAAATGGAGCAGCCTGAGAAACAAACAATGTATGAAATGGGAAA
ACTAAAGGGGCCACTCTTCTCTCTTTTCTGAAAGGCTTGCAGACTTGGTG
GTGCACCTGGAGAGTTTATGGATGGAGATGGAAGCTCTTAGGAGACAAGA
AGCATGGAAAAAGAGAACAAAGGCTCAGTCCCAGTGACTGAAGAGAGCAG
GAGTTTTCCAAAGAAGGTGCATGGGAGGGCCACTGGTCAGAAAAAAAAGG
CTGAAAAATACCAAAGGACAATGTGCTGAAATAGCCCATTTCAAGAGAAA
GGGTTCATCTCAAACCAGCATTCTGACAGAGTGGAAGGAGGGGTGGCTCA
GGGTTATGAGATCACCATCAGCTTTTCCAGTTTTCCCATATAGCATATGC
CTGCTAATGGTATGGAACAAGAGTAAGGCAAAATAGGATGGTGTCCTATA
GAAATGATAGCTCTAAGGTGTTTTTAAAAGGCCTTGATTTCATATGGAAT
GCACTCTCCTTATGTGGAACGGATATTAAATAACCGGGGTACACAAACTA
GAATCCCTTCCCAAGATTGGAAGGGATTGGTAACAGCTGTACTAGAGACT
GTCAGCCGTTGCAATGGTTAACATGCTGGAGGAAAGAAGCTGTGAACATT
GAACAGTAAAACAGAGCAAGGGGTATTAATATAGTGAACGAACAGCTGCT
AGGTGAAGGGCGGTACTCTAGTGTACAAGCACAGACTCGGTGTCATGAAA
CTACTATAGAACAAGGTTGCCTCAGTGGCTATAACACCTTGGGACAAAGG
AGGAGCCAGGAAAAAGTCCAGTTCATTTACAAAGATTATATAAGGCTCTG
GAGAAGCCTTCACTGATTTTTTTTTTTACAAAGATTAGTCTCAGCTATGA
ACAAAGCCATATCAGACCCTGACACAAGGCAGGTGTTGATAGAGACCTTG
GTGTATGACAATGCAAATACCAAATATAAAAATGTCATTAGACTTTTAAA
GGCACAAGTAATGCCTATGGATGAGTGGATAAGGGATAAGACCAATATTA
GTTCTAATGTGTACTGTGCTAATATCATTGATCAAGCTATAGCTAGAGAT
CTCTGATGTCAAAATGCCTTGTGCTTCAGTTGCAGCAAATACAGTAATTT
GCAAGGAGTCATTGTGGCCAAAACTAAAGGTCTCAGATCTCAAAATGCCT
GATGCTTTGTGGGAAATAGGGTCATTTGCAACAAAAATGTGAACAAGACA
TCTTTAAGGGCAATGGTTTTTCTAAATATAAACCAGAAAGACGGCCTAGG
CTTCCAAGGTTGTGCTGGCGATGTGGCCAGGGTTGCCACTGGACCAATGA
GTGTAGGTCCAAAAGAGATATTCAAGGTAACGTATTACCATCATGAAATG
GTCTTGGGGGCCTATCTTGAGGCCCTGCAGCAAAGAGTATGAGCCATTCC
AACCAGAGAGTGGCATGGAGACTCAAAACCTTCACTGGGCACTGGAGATT
TAATGCACACTAGCTATTGCAGGCAGCATGGCTCTAGACTTGGCCACAGA
TAAACATCTTGCTCTATCCCCCAAAATTCAAAGTTATAACATAGCTACTG
GAGTGTATGGTCTTTTTCCCTCAGGGACAGTAAGGATAATCTTGGGAAGG
AGTGGATTGACTTCCTAAGAATTCACTGTGCATCAGGAAGTATAGATGAA
TATTTCAAAGGAGAAATTAAAATTGTGGCATATGTAAAGGTAGAGCTGCA
ACTTAACACAGGCGATAGGGTTGCTCAGCTGCTGCTGTTTCCCTATATCA
AAGGCAAAGCAACTGCAGCAGAAAGAGGAGAGGCCTGAAAACCTTGGGCA
CTGACACAAAAATTGCTTATTTCATTGAAAATGTCTGTTTATAACTTCCC
ACTATACAGCACAACAGGAGGGGGCTTAAAACATAATGGGGAAAATGTCA
CAATTCTGCAATTTTTGTTTCCTTAAAAAAAACACACACACAGAATTTTA
ATAATGTGTTCTCATCTTAATCCCGGGTGTGGGAATTAGGGCTGCTTTGG
ACCATTCCCAGCAGCTGACTATGATTTGCCTCATGCTCTAGCAGAAGTAT
GATTTTTGCCACCTGCAGATAGTTTCTGGGATTGTGTGACATTTGGAATT
TTGGGAACTTTTTCTGAAGGTATATAAATGCTAAGGCCCTGGTGGGGAGGG
TTGGTGGTTGGTGGTCATTCAGGGGGGTGGTTGTGGTTAGTGGTCTTGCT

FIG. 3D (34)

CAAAGAACAAACAAGAAAGTCATTTGATTCAGATGTATCTTTCTTCCTTC
CCCCACTCTTTCTCTCCTCCCCCCGGCACCCTGCCCCCTGCCCCGACCTC
TACCCTTCTTTTTCTATCTAGTGACAAGGATGAAACCAGGGGGATAAAGG
GTGGGAAAAAGAAGAGCCCACAAAGTAACTCAGGTTGGCTACAAGTTCAT
GCCAAGAATCCTAGGACCTTGTTGTTTAAAGGCTTGTTTTATTTTGTGAA
CATGAATGTTAAATGTACATACATGTTAAGTGTATGTATGTACACCATAT
GCATGCATACAGAATCCAGAAGAAAGTACATTATACCCTGGAATGGAACT
TAGAGTTGTGAGACAGCATGAGGATGCTGGGAACTGAACCCAGTTTCTCC
ACAAGAGGAGTAGTTGCTCTTCACTGCTTAACCTTTCCTCCAGCCCCAAT
CCTAGCATTTTGGAGGCTGATGTAGGAAGATTATCCCAAGTGTGAGGTCA
TCTTGGGCTCCATAATAAGTTTAAGACCAATCTCAGCTCCAGAGTAGGAC
CCTGCCTCAAAAACACACAGGTGGAAAGATGGGTCGGCAATGAAGAGCAC
ACACTGTGCCTCCAGGGGACCCAAGCTTGGGTCCAAGCACCCTTGTTGGG
CAGCTCACAACTGCCTGTAACTCCACCTCCAGAGGATCCTAAGCCACCTT
CTGGCTTGGCTTCATGGAGGGAACAGGTATGTGGGTATCTGAGTGTGACG
AATGAGCAGCAAGTGAGTCTCGCTGTGGCTAGCACAAAGTATGGGCTGAA
GAGCAGGAGGACAGCTGAAAAGTGGCCTTTCCTGGTGACTAAGTTGGTCT
GAGCAGCTGAGTCAGTTTCTTCCTGGCTGCTTGGCTGGTCTCAGTGCTTA
TAAGCTGCTCACTTGTAAGTCTTTTCCTAGGAGCCCAGCTTGTCTAGGGG
TTGTCTTTGCAACTGGCCTTGTCTGACAGTGACTTTCAGCAGTCTTAGCT
GCTTATATACACAGTCTTAGGAAAGAAGGCTGGTGAATCTGATCCATTTC
AGGAACTTTCTGAAGCTATTCTGAATTTACTTTACAAGCTTACCTGCAGG
ATAGAGGATCTCAGCTCTTTATAAACATCCTGTCCTAAAACACCCTGTTG
TTCCTCTTCTCTTTTACATCCTGTGTCTTGAGAAGTTTGCCTCCAGGATG
GAAGTTGTTCAATTCAGAGGACACTGTTGCACAAGCTCCCAGCACCCACA
TGTGAGCTCAGTGCTCTCCTTGGCTCTAGCTCTGCCCTATGAGGTTTTTT
ATTTTGTCATCATAATCTTTTCCTATATCCTTCCTTGTTCTGGGAACTCA
TCTGGTTCATTTTTTTGGCATTTTGAGAAAAGCTCTCACTATACAAATCA
GGCTGCCTCCAAATCATCTTTTTGCCTTAACCTCCTCAGTACCAAGATCA
CGAGTGGATCTTAACACTTGACTGACTCGTTTAAGTGTGAGGAAATGTGG
ACCAATAAGAGAGCCCAGGAAAGCCCAGGAGAATCTGTAGCCCCATGGCT
GTTGTGTCAGAACCCAGAGTTTTGTCAACAGAATTTGGTTCCTAATTTCT
CCACTTTATAAAAACGAGTGAGAGAAACAGGAACCTATTCAGATCTGGCG
TCTGAGCAATCAGTGGGTGAACATCTAGAGATCTGTTCTGCATCTCCTCG
CCAGCTGGCAGAGCATGCGTAAGGCGGGAGGGAACAAGGGCAATCACTCA
CTCTGGGGCTCAGGCTTGCCCCTTGGGTCAGGTGTTTCTGAGAGACGTGA
TGTCTGCTTCTCTTGTTACCATCCCTCATCCTCTCCCCTCCTTCTGTCCC
CTACTTACCAATTTCACTGGCCAGTGTCCATATTTCCTGCAAAAGCGATT
TGGTTTAATGAGCTTGACTATGCCCGACTCCTTTAGGGAGGGTGGGGAAA
GGGCAACGAGGGCAGTAAGTGGTTTCCACAACCACTTTGCACCCGGCTGC
TGGGCCCCAAGCCAGAGGAACGTGCATGAGCCATGAAGTTTCCACTGATA
AATCCACAGATGCTTCTAGCACCTGCCTTTCTGACTCAGCCTCACCGTGC
CGCCTGCCAGCTGTGAAATCAGTGCCAACAACAGGTAACCGAGACCCAGG
CGCAGGGCCAGGACAGCTGTCTGACACTTCCAGACAGGATGTGGAGGCTG
ACAGTTGTGATGGAGAGGAGATGGGGAGGACAGAGACGGGCTCAGCTTTA
AGACACCGAGCCACAGAGCACCAAACAAAAGCCAGGGCCTTCTGAGGTAG
AAGTAACAGAAACCAAACAGGCAATTCTACTAGTTTCCTGGGACTGTTTG
CTGCATTTGCCAATCTTGGTAGTTTTAAAAAACAAAAACAGTTTGTTCTC
AGCACTGGCAGAGCTTTCCTCCTCTGGAGGCTCCAGGGGTCCAGACTCTC
CTCTGTGGTACACTGGCTTCAGACATATCTCTTGCCTATGGCTGCCTCAC

FIG. 3D (35)

TCTAAACTCTGCCTGTCCTTGAATTACCTCTCTCTGCACTGGCTTTATAA
AGGAAACATGAGATTGTGTTTAGGGCCTGTTTGGGTGACCTCCTCAGGAT
CTATAACATAATCACATCTCTACCGTATGAAGTGACGCTTCCGTCCCAGT
GTGTAATACATTTGCCGGCGCCTGTCCTTAGGACAGTGACCACCACCAAC
TGTGGAACTTGACTATGTCCACGTCATCTTCCTACTAGCTTTAGAAGGCT
TATACCCACACTTTCTATCCAGAATTGTATTTTTATTTAGAATCATTCCT
ACTTTTAAAAAAGTCTCTGTGGTTAAAAGCATTGCAGAGGGCTTGGGTTT
TGGTCCCCAGGACCCACATCAAGTGGCTCACAGTGTCCTGGAACTCTTGT
TCCAATACCCTCTTCTGGTCTCCATAGGCACTACATACATATGGCACATA
TATGTATACTCAGGCACACGTGTAAATTTTAATGTCTACTTTTTATGCTA
AATATCAAAGTCACTCGAGCAGTGGAGTTGAGCACACTCACATAAGGAAA
TCATCAGACAGACACTTCATCCTGTGTTGGAGCCACTTTGTGGCTGGAGT
AAGCAGGGCAGAGTGATGTTTTCATTACTCTCTGGCCCCAGCACCCCCTG
CCTCTCCCCACCCATTCGTCCATGCAGGTGGGGAAGAGAATTCTCTTTGT
GAAATTGGAAGTTTGGACCCAGCTTCACTCTTACTCTGCCCAGTACCTCC
TGTGAGAAACCCTCCTATCCCAGGTGACCTGCTGGCTGTGACTCTCCTCA
GCAAAAGGCCCGTGACCCACACTGCGCCACTAATGTATCATCCCCAAATG
CTGAAAAGGAAGCGTGTCTTCCTCTCTCTCTCTTTTTCTTTTGGTCTTTT
TGAGACAGAGTTTCTCTGTATAGCCCTGGCTGTCCTGGAACTCACTTTGT
AGACCAGGCTGGCCTCGAACTCAGAACTCCGCCTGCCTCTGCCTCCCGAG
TGCTGGGATTAAAGGCGTGCATCACCACTGCCCGGCTGCGTGTCTTTCTC
TTAGCGGTCTCTGTGGAGATGCTGAGTATGAAGCTCATCCTACCCACCCT
TCAGTGGGGCCTTTTCTAGCTACTGAGCAGCTGTGTGAGGACTCGTGATC
ACAAGGTCCTTTGAACCCTTGAGACAGATGTGCCTGAGCCCAGTTTGACC
TGACAAAAGCCTAGAGCTCACTGATAATGCCAGCAAACACCATCTTTGAG
TTTGCAAAGGAATCGCAACACATGCATTCAGTTTCCGTTGCTGGCTGCTG
CTCCAGAGATGGCTATATTCATTCTCAGGTACTCAGACTCAAGAGTAGTT
CTGGCCACACAGGTCTCCACATTTCGAGGTCAAATGACAGAAAACCAGGT
TGGTCTCAGTGCACATGGGTTTATTGAGCCACTGCAGGTGCTGGGGAAAC
CATGGCAGGGAGATCCTGGGAAGCCAGTGGGGTGCTGAGCAGGAGGGACC
TCAGTCTCTCCTTAATGTCTACACACTGTGTCATAGGTGACAAGCCACGT
CAGTGCTGTGACACGGGTAAGCTTAATGGTGAGTAATGGCTAACTGGGAG
GGTATTTAGGCAGCCTTGTCTGTCAGCCTGTTCATATGATCTCCTTAGTG
CCTTGTCATCTTGGAAAAGGACAGTTCCAAATTCTAGGAGCGGGGGCTAG
TCTCTGTCCTGCTCTGTAAGCCCAGGGGACCCAATGAGGCCTCATCTATG
GGTGCTCAGCTCTAGGATGGGGAAGAAAATGGACAAGATGCCTACTGACG
GGAACACAGGCTTTTCAGTCAGACCCTAGCCTCCAGCCCCCAATCCAGAG
GACAGCCACACAGGGGTCCAGGCCTGCAAAGGGCAGCAGACCTGAGGGCA
AGGGAGTTTCAGCTCAGTGAGCAGTCATCGGGAGACATGGCAGTCAGCTG
TGTCGTCCACGGTTCATGTTCCTAATCAGAGCAGGGCCTGGAGAGCCAGG
GCAGTGAGTGCATACAGCCAGGACACCTTGGGCGTTAGGACAAAACAAGG
ACTGTTTCTGCCTCCAGCTCTTCTCAGGCCACTCGTGCCTTGCCTAGGAA
GGGTAAGAGAGCACAGATGGGAAGGATTCGGAAACTGTCAACTCCCTGTC
CTCTCCCCATACCTACCCGCGGGAAACAGCACCCAGCAGTCTGGTCCTGC
AGAACTGATGGCTGCAAGCTGTCAAAGGCTTGTATGGCACCATCTGCGGA
GTGCAGAGATCCAGAGAAGGCTTGGCCAGGAAACCCTAGAAACTACCCCA
CTCCCTTGGGACAAAAAATAAGACACCCTGGAACCTGCAAGGCATGGCCT
GAGATGGAAGGTCACTGTGCTAAGAATGACCCACAAACTGCTAGTGAGGT
TGACAAGGGCTGCCCCCTCTCCCTTTACAGGTGAACACAATCGGGATTAA
TAAGAGTTTAACTCTCAGCTACTAAGTGGCAGAGACAGGCTTCAAACAGA

FIG. 3D (36)

CCCCCAGAAATCTGGAACTGAGCCATTCCACCCAGAGGCAAGAACAGCAG
AGGTAAGTTGGGCACACATGGAAGAAAGGGCCACCCCATTAGTGTCAAAA
GGGAGGCCAACTTCAGGCCATTGGACACGTTTTAACGCTGACTTCCACCC
ATGTACCATGGCATGTGCACACTGTCCATCGCCCACACCAAACATGATGC
GACGTAAATAAGACCCACGGGCCAGGCAGCTTGGATTGGGCCACAGACAT

FIG. 3D (37)

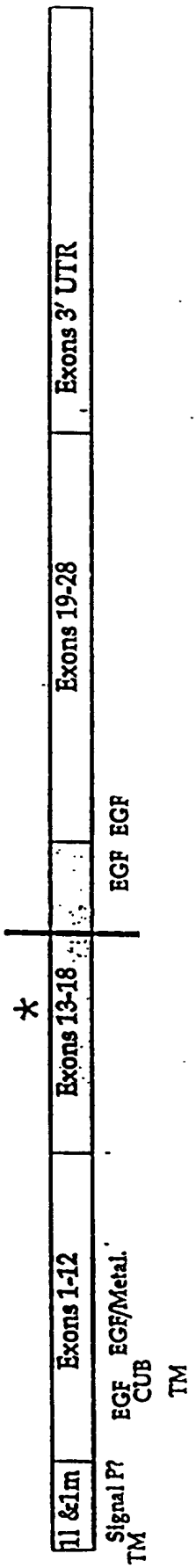


FIG. 4

Exon 1	CelegE106	TCTCCTAGTTGTAGTACATGCTGTTG
	CelegE108	AGGTCCTGTCTCAAGAAATAGCAATAAC
Exon2	CelegE33	TTTGAAGGCCCCTGAAGTCAGAG
	CelegE36	TTGAGTCCCCATCATAAACATATAAATGG
	CelegE37	TTCTAGGCCAAATAGAATAATGAGACTTC
	CelegE40	AGAACTAATTCCATGAGATGAGTGTG
Exon 3	CelegE41	TGAAGTTGCTGTAATCTGGTCTGTG
	CelegE44	AAGGAGCCTGACTAGAAGCCTC
Exon4	CelegE69	TAAACTCCCTACAGTTCACTAACTCAG
	CelegE72	AGCGCTGTTGAGTGTGAATGTTCTG
	CelegE73	AAAGCCACAGTTGTCTGTACAGTGAG
	CelegE76	AGGTCTGCATTAGTTGCAATGTTGC
Exon5	CelegE77	TATACACCCCCTTATATACACTCAG
	CelegE80	AGAGCCTCTCATAAAGCTGTGGTC
	CelegE81	TTGAACATATATCCGCCAACAACCC
	CelegE84	CTTGGAATACTATAAACTTTCAGGCTGC
Exon6	CelegE101	TAAAGCAACAGGAAGAGTTGAACTTCTTG
	CelegE104	TGCACCCTGTGTGCACATGG
Exon7	CelegE109	TTACGGTGTCCTAATAATAAGGGCAG
	CelegE111	AATCATGGGTATTGTTAACTCCGAAAGC
	CelegE114	TGTAACAATGTGTGCCGAGTGTCC
	CelegE116	TCTCTCTCCAGCCCTAGAGTTG
Exon8	CelegE86	AGAAGAGGAGCCTGCAACATTGAC
	CelegE88	TTTGTTGGCGCTGAAAGCCTTG
	CelegE89	TGGCCACAGTAGTGTTTATGATGAC
	CelegE91	TTAATCAATTGCCTCTGCAGATTCTAG
Exon9	CelegE93	TGGCTTACGTATAGGGGGAAATCAAG
	CelegE95	TTGTGTGTGTTCCCTCCAAACACC

FIG. 5(1)

	CelegE98	GGACCATTCTTAAGGACAGCCGAT
	CelegE100	ACATAGTGATCTTTCCATCAGCAAAG
Exon10	CelegE117	TGAATGCACAGAGACCCTCCTG
	CelegE120	CCTCTTACCATTCAGATACTGTTAGG
Exon11	CelegE121	AGCAACAACTCAAACCAGCCCTAC
	CelegE124	TTCTTCAGTTGCCAACTCCCAGG
	CelegE125	AAGCTGCTTGTGTGGCAGCAG
	CelegE128	AGTAAGGTGAACAGGAAAGTACAGAG
Exon12	CelegE130	TACATAAGAGAGGCTGCCGCATAG
	CelegE132	CCCTACACTCACACTCATCTAGC

Exon13	CelegE30	CCCTGTGTTCCAGATCTCCATTG
	CelegE32	TTCCTAGGTCCACCTTGATCTGAG
Exon14	CelegE14	AGCACCTGAATTCAAATCAGGATGAG
	CelegE15	AAACCAAAGTTCTGAACACATTAACTCAC
Exon15	CelegE17	CTGGTTGCATTCATAGCTGTGTTTC
	CelegE20	ACAGAAGCCAGCATCACTGGG
	CelegE21	TTACTGGTGCTGGGAGGATATGTC
	CelegE24	ATAAGTACTTCATCACCTCAGCGCTC
Exon16	CelegE1	TTGATCTTAGCTGACCAGTGTCTC
	CelegE4	TCTGCATGGACTTGAGCAGAAAGTC
Exon17	CelegE6	CAAATCTTGTGATAGTGAATTACAAGTTGG
	CelegE8	TTTATAGCTGCCCTCAATACATTTTCC
	~~Celeg~~E9	TGTACCTGCAGCCATTGCTTGG
	CelegE12	GGATCTGGGCTCTAGTTTATGTACG
Exon18	CelegE25	TTGAACTATAGGCACAGACAGCTG
	CelegE27	AACTTGACCTGTGTGACTTACGC
Exon19	CelegE193	TCACAGTCTATGGTAATCTGTCAAGC
	CelegE194	AAGGGCAACAATGCCCTGGCAA

FIG. 5(2)

Exon20	CelegE195	TTCCTGCAAATGGGATAGTCTCTCTG
	CelegE196	ATCCCCCAAGCATTTATCATTCTCAG
Exon21	CelegE197	TGTGTTTCCAGAAACCTGCTTTAGTTTG
	CelegE198	TAGTACTTTTGTCCAGGATGACCAAG
Exon22	CelegE199	TGACAAGAAATGTCATGTCTTAACATAAGC
	CelegE200	TTCAGAGCCTCCTTCCCCAACT
Exon23		
Exon24	CelegE203	TAGTCTGTAGCTGAGGCCATTTTGC
	CelegE204	AAGCAAGCTGCAGTTAAGGGACTGT
Exon25	CelegE205	TTGGGACCTTGAGGATTGTTCCC
	CelegE206	CACTCAACAGGTAAAAGTGATCTGCC
Exon26	CelegE207	TGCATCTGATCAGTTTGAATCAGAGAG
	CelegE208	AAACTGAGGCCTGAGTTCTGAAAAGC
Exon27	CelegE181	CACCAAAGCTCTGTACCACTAAGC
	CelegE182	TGACTGTGCAGTGATGCAGGG
Exon28	CelegE171	TTGACCTTGACATTTAGAATAGCCCTC
'UTR?	CelegE172	GCTGAGAATTAATAATGGTCTTTCTCTTTG
	CelegE173	TACACAGTGAGACCCTGTCTCC
	CelegE174	TAGCTGAGGTCCCTTGTGGAAG
	CelegE175	AGTGTCAGAGGACCATGCTGG
	C.elegE176	CTTGAAGCGTCCAACTCATGTGC
	CelegE161	AACTCATACATTTGAGCACTGTTGCC
	CelegE162	TGAGGAGGTGGGGAATGCTAATG
	CelegE163	ACATAGCAGAGGGCTGCTCAC
	CelegE164	ACTGACCTGTGGGAACCTGTG
	CelegE165	AATGCTAGGCATCATCCCTCTTCTAG
	CelegE166	AACATCTAATAGGGACTGAGTGACCC
	CelegE167	TTCTGTGGTGCCTTGGCAAGAG
	CelegE168	CACACATACACACACACTAAGTGCC

FIG. 5(3)

CelegE169	TGGTAGTTACTTAAAGTTTACAAGAATGTAGG
CelegE170	AAATGCTGGGATAAAAAGCATGAACCAC
C.elegE145	TTCAGTTACCTAATGGGCACAAGGC
CelegE148	ACGACACTGACCTCTTGCACTTG
CelegE150	TGTACACCCTGAATGTCTGAACATTC
CelegE152	GCGTTCATTTCTGCTCCTGTAATGG
CelegE153	TGAGCTCTTAATCCCTGCCATTCC
CelegE154	TAGGGCTTGTCCGTCCATAAGG
CelegE157	TGTTACGGAGATGAAAGGCTAGACC
CelegE158	TAAGCCCTGCACTAACGCACTC
CelegE159	TGTTTTGAGAGGTGAGCCTTCTAGC
CelegE160	CATGTCCTACAGTTCTGCTATCACC
CelegE141	CTTTTCTTCATCCAATTCCCCACGAG
CelegE144	TCTCTAAGCTGCACTGTTGTGGCT
C.elegE129	TGGAAGCCAAGAGTCTTGAGTTGC
CelegE132	GTCTGCATTTTAAATGCAGATGTGAAGC
CelegE134	CGAAACGCACGCACATTTTTACCAG
CelegE136	GTGTGATTTAGCATCTGTCGCACTTG
CelegE137	TGTATGTATAACCCAACAATCGCTGC
CelegE140	TCCAGAGTACAGGGAGAAACTAAAGG

Fig. 5(4)

AGCGCTATTCAGCTGTGCCTCCTTTGCTGTCTTGGCTCCTCCTGGAGCACTAT
ATGCACCCATGTCCTTACCAGGCCTTTCACAGAC**GCTGC**CATTGAGAGGGT
TGATGCAGGTTGCAGCCTTTAATCCCCGAGTACTAGGCTCTGACAAGATCCCA
CAGAAGCCAGCATCACTGGGCTCAGATGGCATCCACTGCAGCAAACTATTTG
TGAATGGAGACATATCC

FIG. 6

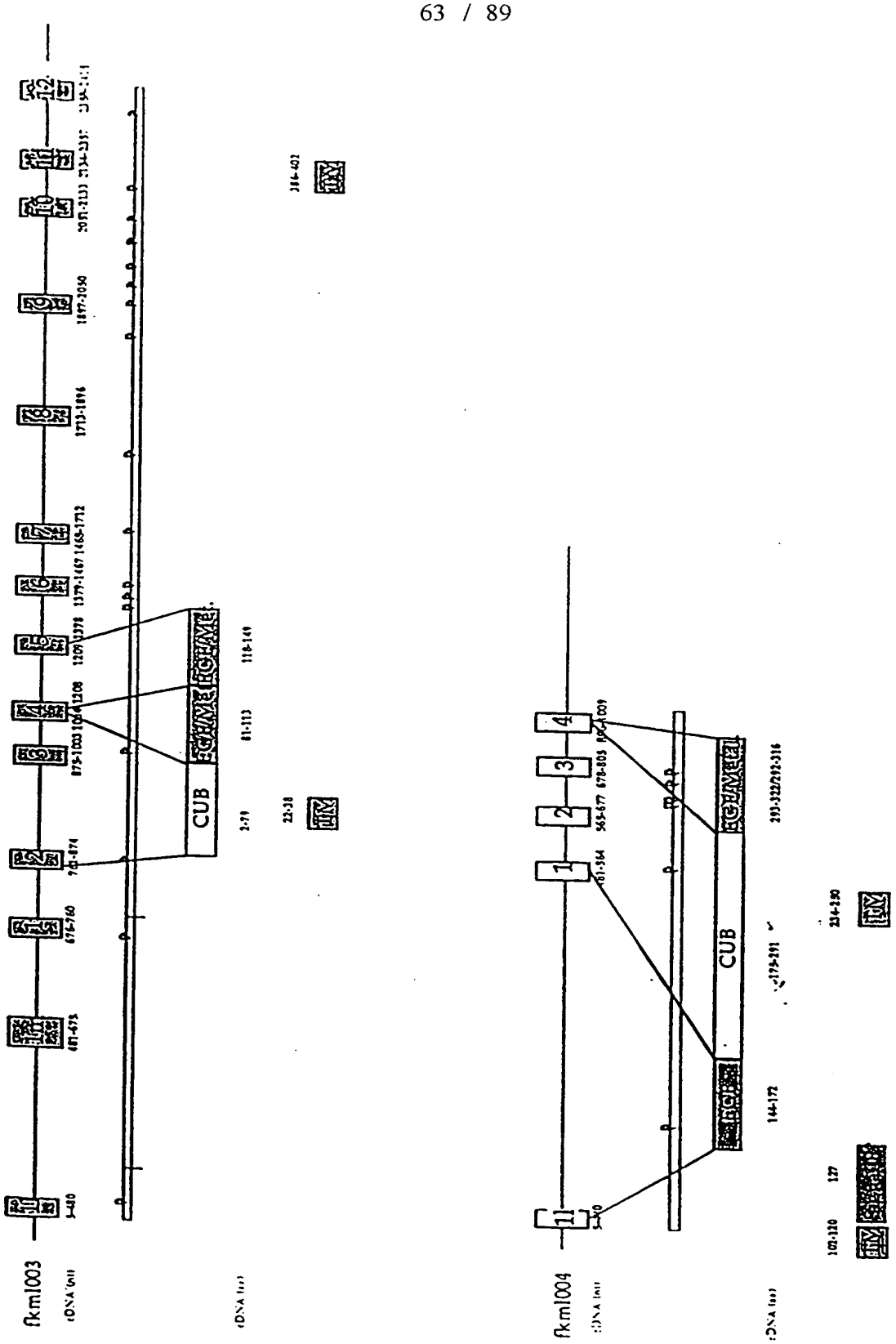


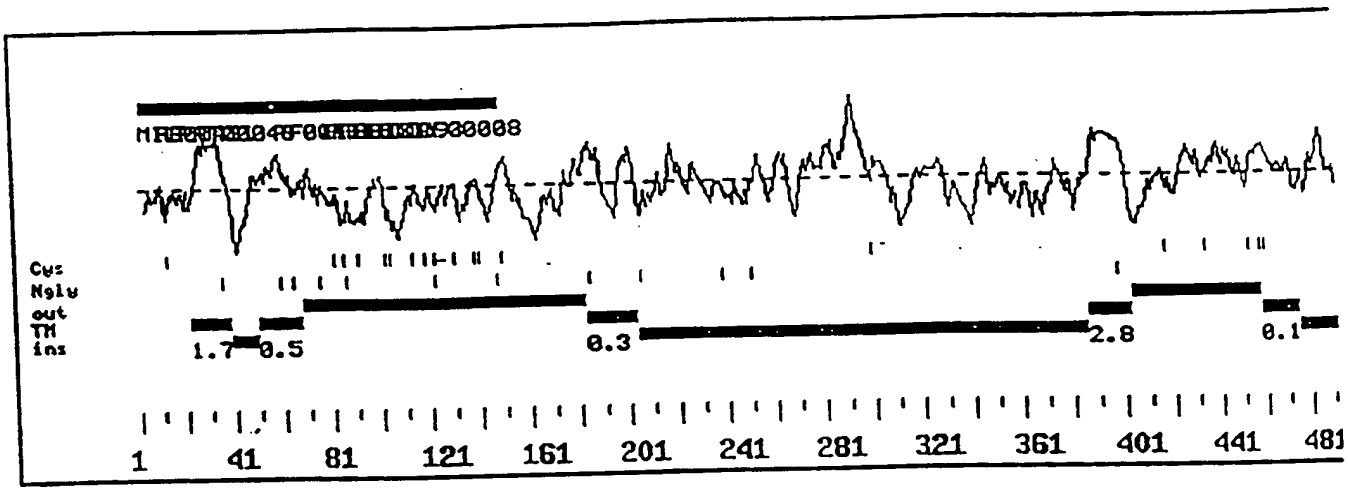
FIG. 7

GAATTCCGGGCGAAGGGGAGCCGGCGTGCGGGGTGTGTATGTGTTCGCTGGGCGCCGGCTCAGCCCCAGGAAGATGGTG
GCGGTGGCGGCGGCGGCGGCGACTGAGGCGCGGCTGAGGGGGAGCACGAGGACGACAGCAGCGCCTGCGGGCAGGAAGG
GCAGGCAGCACCGACCCTGCACCGCGACAGGGGCCTGGAGGCCGGGACCGCGCGCCCGGCTGTGTCTCCCGCGGGTGCT
GTCGCGGGCGCTGCCCCCGCCGCCGCTGCTGCCGCTGCTCTTTTCGCTGCTGCTGCTGCCGCTGCCCCGGGAGGCCGAG
GCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCG
GCCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCG
CTTCAGGACATCTGTCTCACGCCTATAATCACAGCTGTTCGGAAGGTGAGGCTGGAGGAACAGTTCGAGGCAAGCTTCG
GCTACAGAATAAGTTCAAGAGTAACCTGGGGCAACTTGGGCTTGTCTCCAAAACCAAAATGAGCGAAAAGGAGCAAGCT
AGAGTCTTTTGGGAAAATTTTAGCTGACTAATTTTTCACCGAGAACTAACTGGCTCTTCTGGATTTGTAACAGATGGAC
CTGGGAATTATAAATATAAGACGAAGTGCACATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGCTTCAA
CCATTTTGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCACCTCTGATTGCTGCC
TTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCCTGAGGTCACTGTCACTTCAGGTTATGCACTGC
TGCATTTTTTCAGTGATGCTGCTTATAATCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTG
CTCAGGCCGAGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTGGAAAGGGGAG
TCGTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGGCATCTGTAATGCAAGCGATACCAGAGGGT
GCTCCTGCTTTCCTCACTGGCAGGGTCCTGGATGTTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGA
ATATTCTGATTTAAAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGGCGGATAT
ATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGACTTCTAGGGAATGGCTTCCACTAAACCATTCTG
TGAACAGTGTGGTTGTAAGATATGGTCATTCTTTGGCATTACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGA
TTCAACAGGGAACGTGACCAATGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACTCCGAAAGCT
AAGGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGGCATCTGGCCGTGTGGTCATGTTGGTCATCT
TCGGTCATTGCCCACTCTATGGATATATAAGCGTTGTGCAGGAATATGACTTGGAAAAGAACACATGGAGTATATTACA
TACTCAGGGTGCTCTTGTGCAAGGGGGTTATGGCCACAGTAGTGTTTATGATGACAGGACCAAGGCTCTGTACGTTCAT
GGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGATGACCTCTACAGATACGATGTGGATACTCAGATGT
GGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCATACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGG
AGGGAACACACACAATGACACTTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTTCATGGCTTATGACATTGCT
TGTGACCGATGGTCAGTGCTTCCCAGACCTGAGCTCCATCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTGTACA
ACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCCTCAGTGACGTCTTGGTCTTTACCTCGGAGCAGTGCGA
TGCACACCGCAGTGAAGCTGCTTGTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGT
ACCTCCTGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAGTGTTTTTCTAAAAGAACCCTTGACCATG
ACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATACCAA

FIG. 8A

MRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNETAPEVTVTSGYALLHFFSDAAYNLTGFNITYNFD
MCPNNCSGRGECKSSNSSSAVECECSENWKGESCDIPHCTDNCGFPHRGICNASDTRGCSCFPHWQGPGCSIPVPANQS
FWTREEYSDLKLPRASHKAVVNGNIMWVVGGYMFNHSDYSMVLAYDLTSREWLPLNHSVNSVVVRYGHSLALHKDKIYM
YGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKDQYAVVGHSAHIVTLASGRVVMLVIFGHCPLYGYISVVQEYDLEKN
TWSILHTQGALVQGGYGHSSVYDDRTKALYVHGGYKAFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSG
TMLVFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPELHHDVNRFGHSAVLYNSTMYVFGGFNSLLLSDVLVF
TSEQCDAHRSEAACVAAGPGIRCLWDTQSSRCTSWELATEEQAEKLKSECFSKRTLDHDRCDQHTDCYSCTANTX

FIG. 8B



Transmembrane Segments Predicted by MEMSAT

Start	End	Orient	Score
22	38	out-->ins	1.7
50	67	ins-->out	0.5
183	203	out-->ins	0.3
386	402	ins-->out	2.8
458	474	out-->ins	0.1

Signal Peptide Predictions

Method	Predict	Score	Mat@
SignalP (eukaryote)	NO		

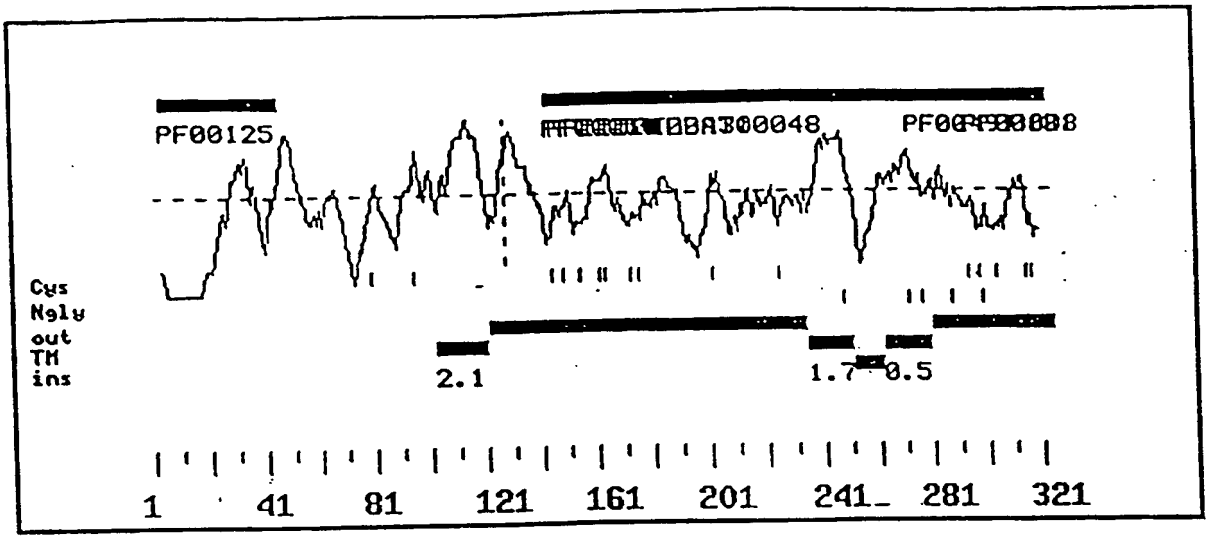
FIG. 8C

GAATTCCGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGAGGGCGAAGGGGAGCCGG
CGTGCGGGGTGTGTATGTGTTCGCTGGGCGCCGGCTCAGCCCCAGGAAGATGGTGGCGGTGGCGGCGGCGGCGACT
GAGGCGCGGCTGAGGGGGAGCACGACGACGACAGCAGCGCCTGCGGGCAGGAAGGGCAGGCAGCACCGACCCTGCACCG
CGACAGGGGCCTGGAGGCCGGGACCGCGCGCCCGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTGCCCCCGCCGCC
GCTGCTGCCGCTGCTCTTTTCGCTGCTGCTGCTGCCGCTGCCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTG
TCCGGCTCGGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGCC
AGTGCGTCTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGCTCTTCTGG
ATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACGAAGTGCACATGGCTCATTGAAGGACAGCCAAATAGAATA
ATGAGACTTCGCTTCAACCATTTTGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACG
CACCTCTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCCTGAGGTCACTGTCAC
TTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTATAATCTGACTGGATTTAATATCACTTACAATTTTGAC
ATGTGTCCGAATAATTGCTCAGGCCGAGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTG
AAAACTGGAAAGGGGCCGGAATTC

FIG. 9A

EFRKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKGGRRGAGVRGVYVFAGRRLSPRKMVAVAAAAAT
EARLRGSTTTTAAPAGRKGRQHRPCTATGAWRPGPRARLCLPRVLSRALPPPPLLPLLFSLLLLPLPREAEAAAVAAAV
SGSAAAEAKECDRPCVNGGRCNPGTGQCVCPTGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTWLIEGQPNRI
MRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNETAPEVTVTSGYALLHFFSDAAYNLTGFNITYNFD
MCPNNCSGRGECKSSNSSSAVECECSENWKGAGIX

FIG. 9B



Signal Peptide Predictions

Method	Predict	Score	Mat@
SignalP (eukaryote)	MAYBE		127

Transmembrane Segments Predicted by MEMSAT

Start	End	Orient	Score
102	120	ins-->out	2.1
234	250	out-->ins	1.7
262	279	ins-->out	0.5

FIG. 9C

ATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG
AGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGC
CAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAG
GTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCG
GGCCTTCGGGAAGGTYCAATGTGTCCTACTKGGTGCTGGGGAAGGATATGKTCCCCATTTTACAAATAGTTTTACTACA
GTGGGATGSCCGTCTTGAGGCCCAGTGTTGCTTGGRATTCTGTGGGAATTTT:ATTCAGGAACCCAGTTACTTCGGGGA
CTGAAGGCTGCAACCTGCATTCAACCCACTYMAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGGCAGT
GCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTG:CATGTGGTGCAGCAACA
TGAAG:CAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGG:CCAGTGTATGGAATGGTATACGATGAGC
ACCTGCCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTG
ATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTAC
AGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCAC
TGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAG
GCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAA
TGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAG
CTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGT
TCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAG
GGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACC
CAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTG
ATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGATTGCCTT
CTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTG
GCTGCTGTGGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGA
TGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTAT
AAAGACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTC
CCTCGAGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGC
AGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTG
CATCTGATGCTGGGGCCAGGGACTCTCCCACGCACGAGCTAGTGAGTGGCACACCAGAGCCATCTGCAGGGAAGGGCGT
GGCGGGGAAATGGCTGTGCGGTGCGGGACGGAAGACTGGAAACCCTCAAAGCATCTGACTCACCTGCATGATCACAAGC
TTTCTTTGACGGTTTCTCCCATCCGTGTTCCAGCATCTAACCTTTTACTTTTGCATAGGAAATACTTGATTTAATTACA
GGTCCAGGGATGAGCTGATGGTTGCTGGAGGAGGCCAGTGTAGAGCCAGTGAGAGAACTAGGAATGACACTCAGGTTCA
CTGTGGAAAACTGTTCTTGGGACTGTCTCAACTGTGCAAAAAACAAAAGATGGAGTGTTTACAAGTAGACATTCGTCAT
CAGTTGTTCTTGAACATGGTCTTTTAAAAACTAGTCAGATGAATTAACTTGTTTTCATCTGAAGCCTGCTATCTTTTTT
AAAAGATGTGCTATTTATTCTTGCACGATTTAGGCAATTATCTCTCTTCCAGGGAGTACCTTTTTTTCTAGTTGAGAAT
TAATAATGGTCCATCTCTTTTGATCATATCAAGCTAGGATAGAAGGGGGGCTATTTTAAATGTCAAGGTCAGCAGTGTT
ACTTTGAATGTAAACTGGTATAATAGGTAGTTTTCTATAGTAACTTGATTAATTTAGTCTTAATCCATTTGAAACTCTC
TCTTCCTTTCTCTCTGCCTGTCCCTCTCCTTCTCCATCTCACCCTCCCTCTCTCACACATACACACACAAACACATACA
CACAACACTAAGTGCCTAGACTTTAAATAGATCTAGCAATTGGAAAGTTAGTAAGCCTAAGTTTTTACATAATTGCATT

FIG. 10A (1)

CCTACATTCTTGTAAAATTTAAATAGCTACCATTGGCAATCTGCTTTTTTTCTAAAATCTGATTTGCAGCCAGGAAAGA
ATTTTCTCACCCAAGGAACATTTGATCTAGCAGCAGGGATGAGAGGAAAGCAGAAATGAATGAACTGTGAAAGCTCCTG
TTTTTATTATCAAAAAGGACACTGTCAAGAAGGCGCCCCCTGCCCCCACCCCCGTGTCACCCTAGGCCTGATAAGCGAT
CAGAGGAAAGGACTCATTCATGTCACGCTTCCTTGAGCAGAAAAGAGCACTGAGAGCACTTGGGACCCCTGGATCAGAG
AGCATCTGTGTGTCCTGCAGCCTCCTCTGAACTTGTGGTTCATTCTCAGGCTGGGGTGGACTCAGATGCCAGGAAAGGG
ACAGCCTCCCATTGTCAGGCAGAAGCTGCCCAAAGCCTGGAGAAGGACTTGTTTGCCCTCTTTCCCCCAGGAGGGGCTC
GACCCACCCACCCTCCCTCTCAGACCAAGGTGGTGGCTGTGAGGAGGGCAGCAAATGCTGACAAGGATGAAAAGCACAT
GGAAAAAAATGGACGAGGAGGGAAAACTCTGCCAAATGGAAAATGACCAAATTTAAGAGGGTGGGACAGTCCCCTGCTC
CTCTCCCAGAGGGCACTGCTTGGAAATTGTGTTTTCCCCATTTATGGTGCTCTGTATTCTGGCATTATGCAGCAGCCTC
CCAGAAGCTCTCTTCTGCTTCAAAACCTGGGATCTCTGGCATTACCCTATTGGGATGGACCGCTGGACAGCAATGCTCG
AGTTTGTGAATTTGGAGAGATACTCAAAAGAGCTAAAACTGCAGCATTTTACCTTTAAATGCAGTGCCTAGAGAGAGAG
TATTGTCTCTTCCCCAACACTAACCCCACTCCCATGAAGAATTGCCTGGAAAGATGTTTTCAAGGAATTTGAACCATAA
AACACTATCTGATGCACAGAACACCTCTACTTTGAGACTCACCTCTCATAAAGCTTCTTTTTCACATTACTGTTAAAGA
CCAGACGTTCTAGAAAAGACCCCTCCTCTCATGAGCTCCCCCATCCCTGCTACAGAACACAGCACCCATGGCGCCTGCA
GTGGACTGGCCCCTTAATTCCCACAGGCCCCCCCAGCAAGGCCAAAGGGAGGCCCCTGGGTATTGTCCTCCTACAAGGA
AGATCCTCTTTGTTTGTTCAAAGGACCAGTTTTCCTAGGCCAAAGAAGTCTCTTCCCCATGTTAGTCCTATGCCTTGAA
ATATCATGCACCATGACCCACAGCCATCTGGTTATGTCTTATTTTTTTCCTAAAAGATAATGTTTATTTTTAAAAAGGA
AGGAAGAAGCAAGTGAAGTTTCATTCTGCTCCAGCGGTGGGGAAGCCGCTGAATCCACCTGCTTCTCCTTTGCAACCGA
CAGCAAACAGCTTTCTCCGGCCTCAGGGCAGAAAAAGGGAATGGCAGGGAGTAAGAGGCGCTGGGCTCGGAGCCTGTTT
CCAAGAAGGAATTGGTTGTCATCTGGCAGTGTTGCGCGTCACAAGAGAGCCTGTATATAAATTAAAATAGTCAAGACAA
CACTGACCTTGCACTTGTACATAACTATACAGTAGTGTCCAGAATGTTCAGACATTCGGAGTGTACATAAAACAGAAAA
AATCTTCATGTATTTTTATTAAATATAACAATGTCTGAGTTTCACCTAAGATGTTTTTGTGCCATATGCTGGATATCCA
GGTTCTCGCCAGGCCCCGATACATGAATAACAAACCCAAGAAACGCATCCCCATTGTGTGATGTGTTCAGATGCATCTG
GCACCAATTAGGTATTTCTTAAAACAGGACTCATCTGTCAGAGTGCACATGAAAAATCAGGCAGGGAATCGAAACGACA
GCGCTGGAGGAGACTCAGGAAGCAGAGGCGTCCCTGCCGCTGCCCTTGGCCCTGCAAGCACATCATGACCCTTTCTGGC
AGCCTCTTGGTGCTCTGGGTAGTGAGGGATGACCAGTCTTGTCCTGAGAAATGTTTCTCTTAGTCTTTAAGTTCAAAGA
CTAACCTGTAGCAATCAGACTTTCCAAAAGGGGGTTCTCCATTTTTTGTAGTTTTGTCTAAATTTTTAATGACCATTTC
CTGGAATCAGTTTATTATACTGAAAACTGGGGGTGGGAGTAGGGAGCTAGTTTGTTGATAAATAGTTCCCATTTCCCCG
TGGAGAATTTGACATACCCTGGACTCCTGTGTGCCTCCTGCCATCCCTGCACACAGCCTGGGGAGAAGCCTGTGCCTCC
CCGTGTGGAGAGAAGGCAACCCCAGATCCCCTGAGCTAACCCGGAGGAAAGGCAGTCCTGGACAGAAGACTGTCAGCAG
AAGGAAAGTACTGGACTACCCGTGGGTAAGTCCTGCCATTCAAGACTGGAGACACCTGGGAAATAAAAAGAGCAGGGCA
CTGCTGGTGGGAAGAGGCATTTTACCTTCCAGTGCAAATCCTGCTCCTTTGATTTAATGGGGTGTACTGGGGCCAGGGG
CTGATTCACTTCCTTGGGAGATGGTGGTGTTTTCATGAACATCTTTGATCCTTCCATTTCATTTATTCATCCATCCATT
CAACAAGTATTTGCTAAACACTAACTTAAGCTAATGCTAGGGTAGTGACTGAGATGTAAAAATAGATTTTAGAATTAAA
ACAAAATCCAAGTCCTCACACCCCTGTCATCCCAGGAGATCTTTCCTTGTGGTGGTTTCTGTGAGAATTGGCCATCCTG
AGGACACAGCCAGGACGGCAGAGGCCTCCTGGCCTCAGGGCATGCCCTGCCTACCTTCTGAAATGTTTACCCCATTGAC
CAAACTTGGCTCCAGCCATTGCGGTGGTTTCTAGATAGCCAGGCCCACCAAGAGATATTGCCCCTTGATGAGAGTCAAA
CACCCTGCCTACAAGGAGATGTTTTGAAATGGAGAGGAAAATTGGCACCTCATCTTTTAAAGGCAGTAATGGAATTGAT
TTTCAGTAACTGAATTTGTGCACAAAACATTCTAAACACTAGTGAAGCCTGTTTCGTTGAACTAATTCTGGCTCTGGAA
ATGTTTTTGTTTTATAGTTATTTACGATTTCGTTTGTTTGGATTCAAGCTTAGTTTGTTAATATGTATAATTTAGCATC

FIG. 10A(2)

TATTACACTCATGTAAATATGGAGTAAGTATTGTAAACTATTTCATTGCGGGGATTGTGGGTGTTATACATACATTTAG
GACTGCAATTTTTTGGTATTTTTTGTATTGTAAAATAACAGCTAATTTAAGCAGGAACAAGAGAACTAAGGGAGGTCTG
TGCATTTTAAACACAAATGTGAAGAACTTGTATATAAACAAAAGTAAATACTATAATACAAACTTCCTTCTGAAATAAA
AGTAGATCTGGTAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAGGGCGGCCGC

FIG. 10A(3)

MYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLVGNSCLKITTAKENYDNAKLFCRNHNALLASLTTQKK
VEFVLKQLRIMQSSQSMSKLTLTPWVGPSGRXNVSYXVLGKDMXPILQIVLLQWDXRLEAQCCLXFCGNFXSGTQLLRG
LKAATCIQPTXMVVSVKGLQTTVLRQCRTPCALRTACGDCTSGSSEXHVVQQHEXSVWTPMPMWPPSLLXQCMEWYTMS
TCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQAPTGNFYPQPLLNSSMCLEDSRYNWSFIH
CPACQCNGHSKCINQSICEKCENLTTGKHCETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQ
LCEVENRYQGNPLRGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKWFNLNITWAASFSAGT
QAGEEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPIKIQIAFSQHSNFMDLVQFFVTFFSCFLSLLLV
AAVVWKIKQSCWASRRREQLLREMQQMASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRL
PRGLGGIPPPGQSGLAVASALVDISQQMPIVYKEKSGAVRNRKQQPPAQPGTCI

FIG. 10B

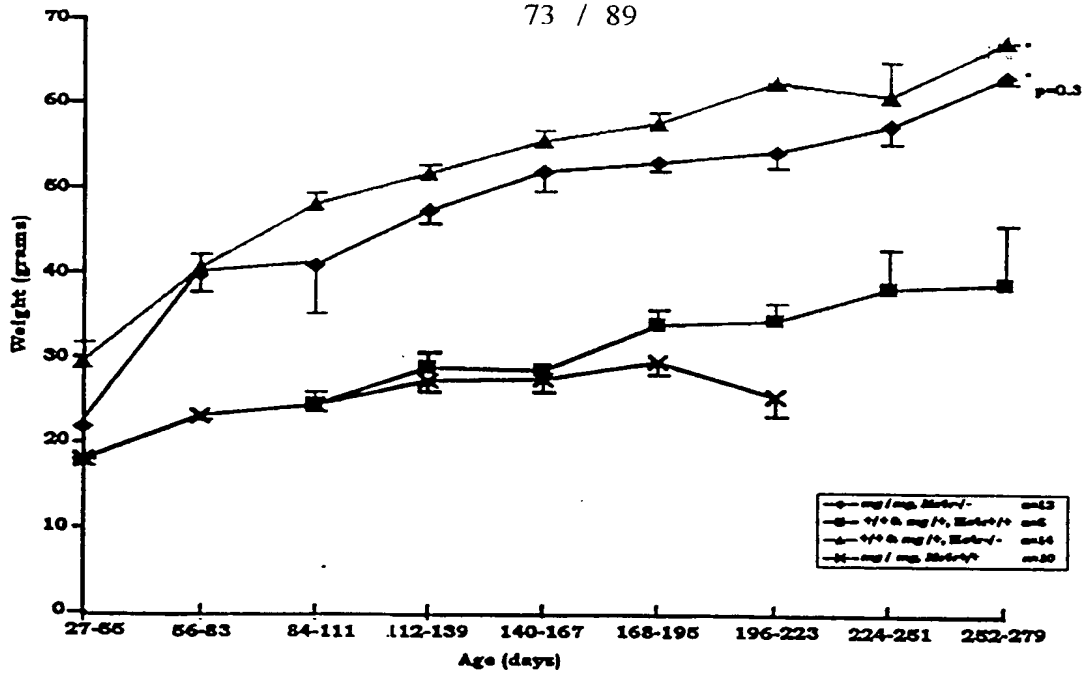


FIG. 11A

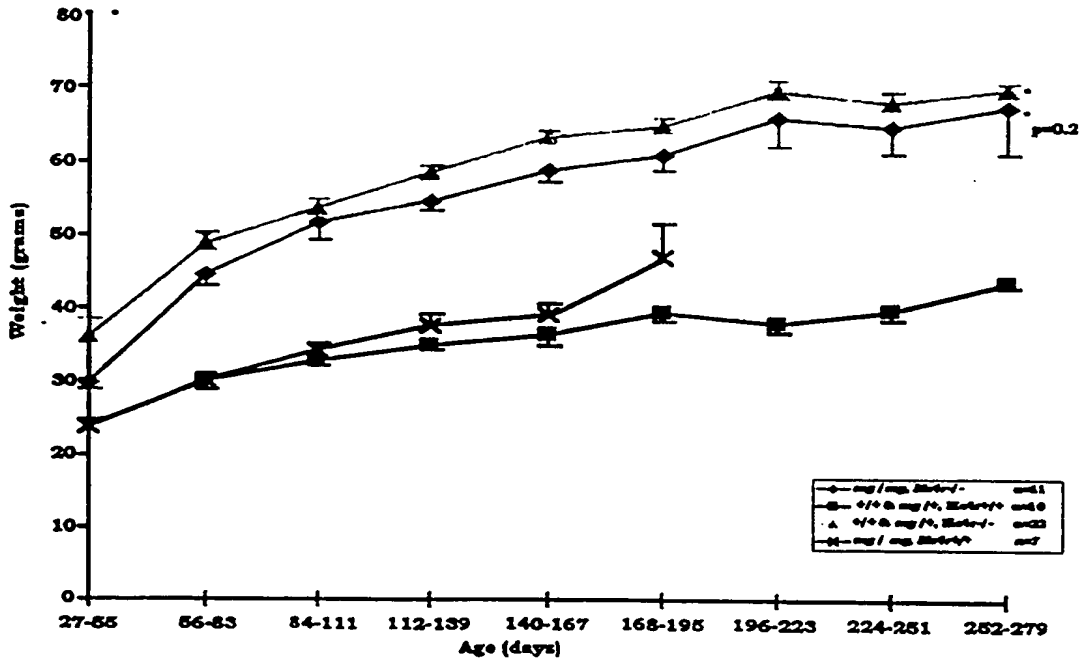


FIG. 11B

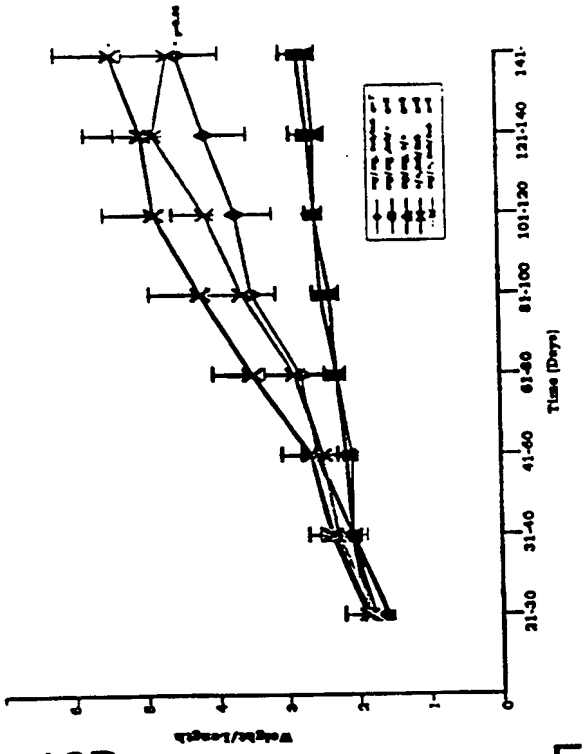


FIG. 12B

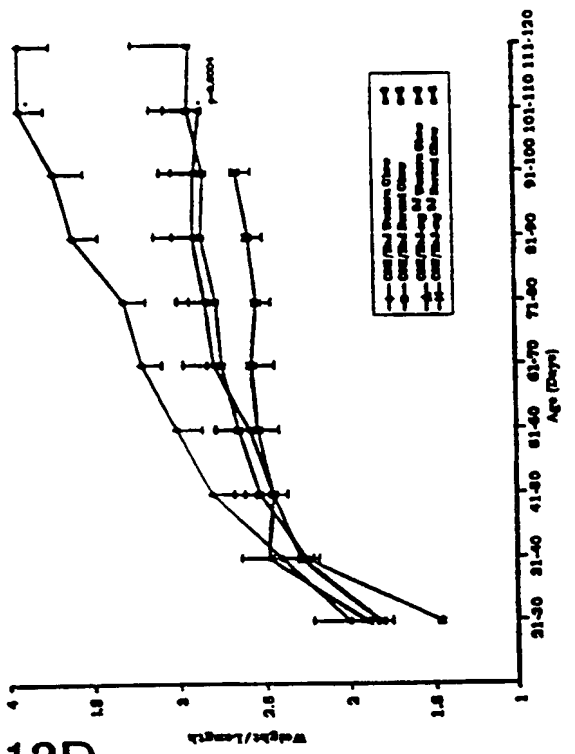


FIG. 12D

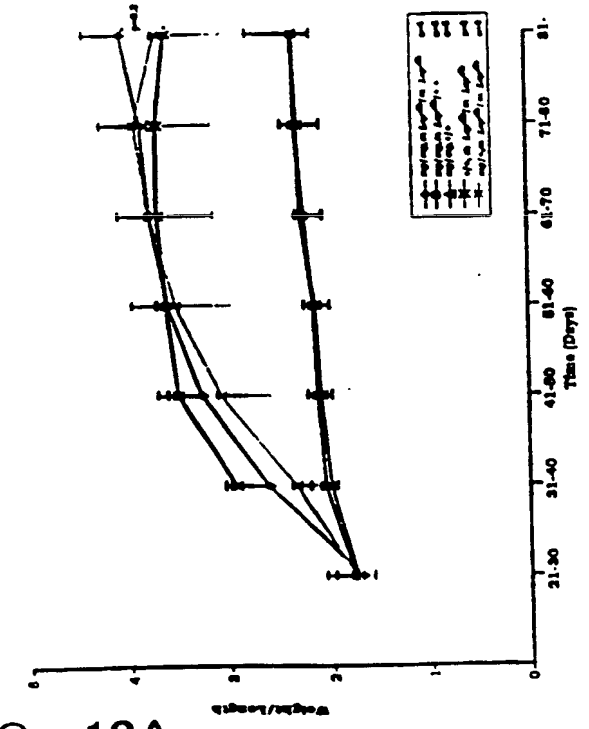


FIG. 12A

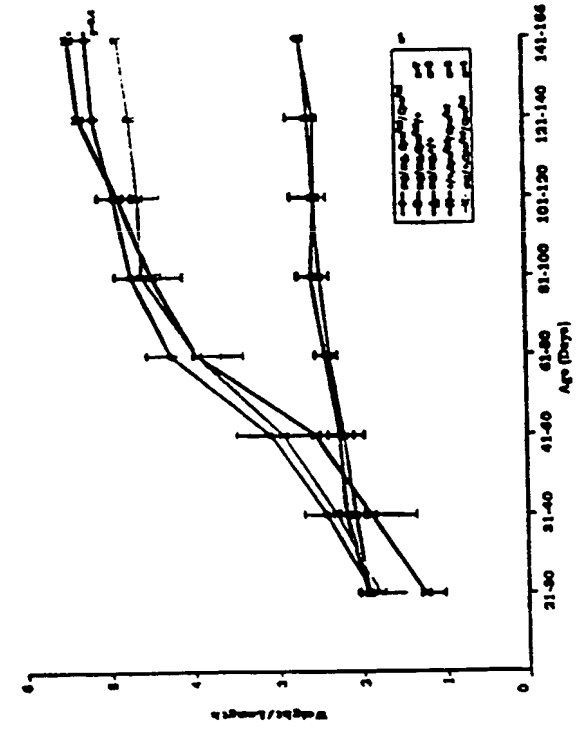


FIG. 12C

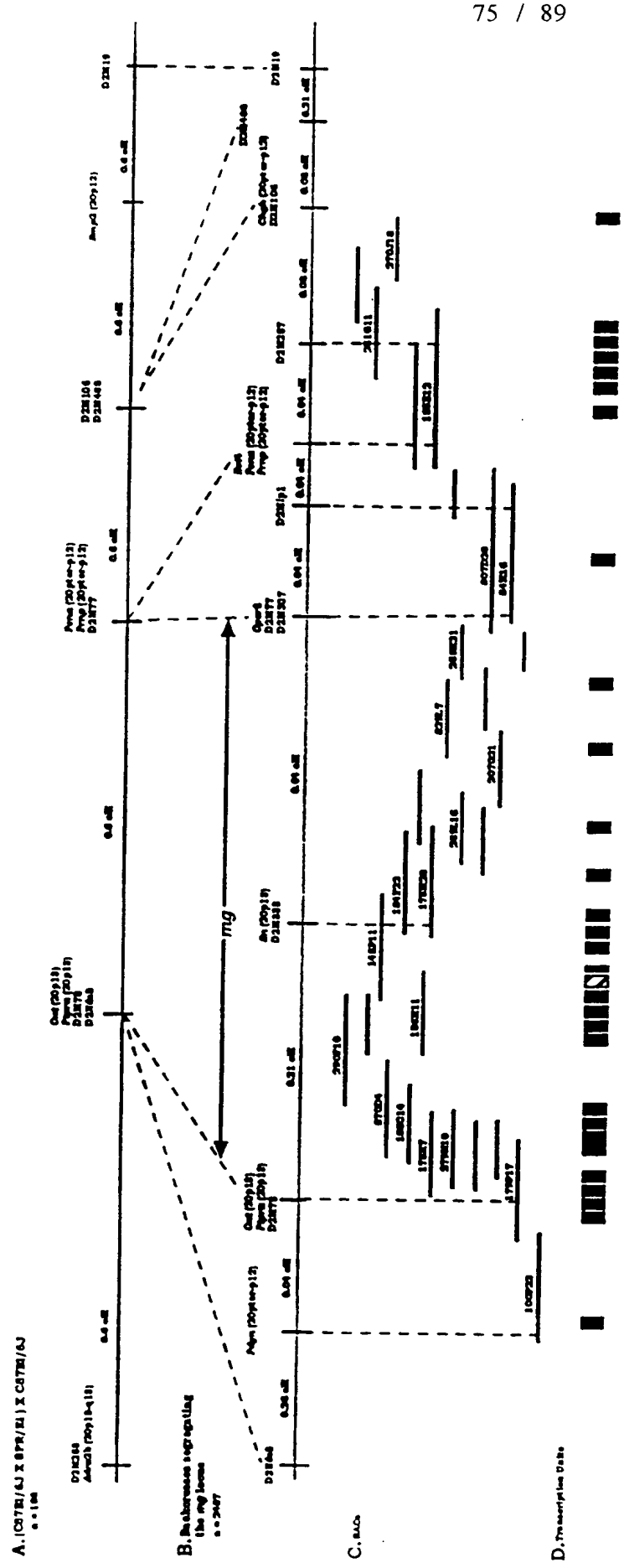


FIG. 13

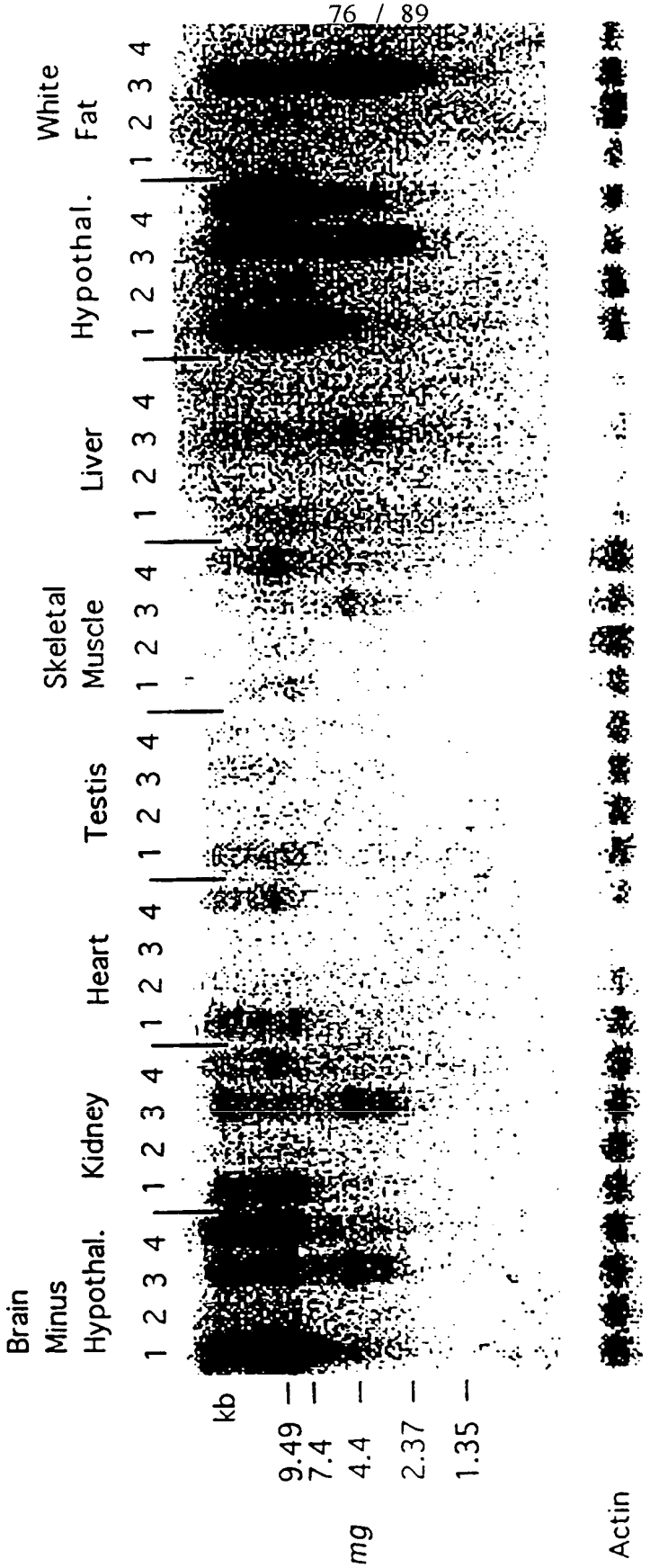

Brain Minus Hypothal.
Kidney
Heart
Testis
Skeletal Muscle
Liver
Hypothal.
White Fat
1 2 3 4
kb
9.49
7.4
4.4
2.37
1.35
mg
Actin

FIG. 14

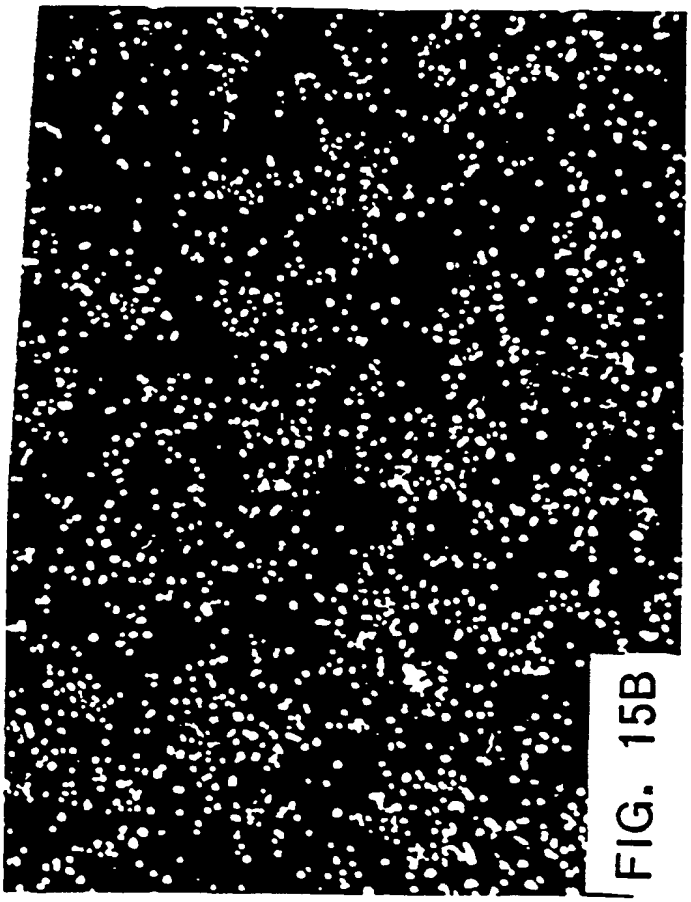


FIG. 15B

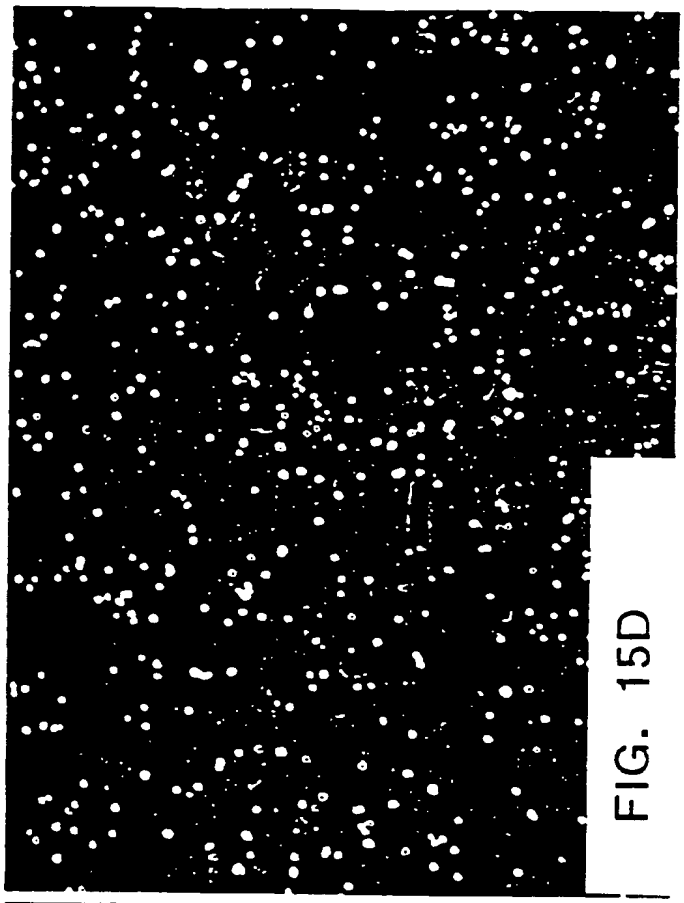


FIG. 15D

FIG. 15A

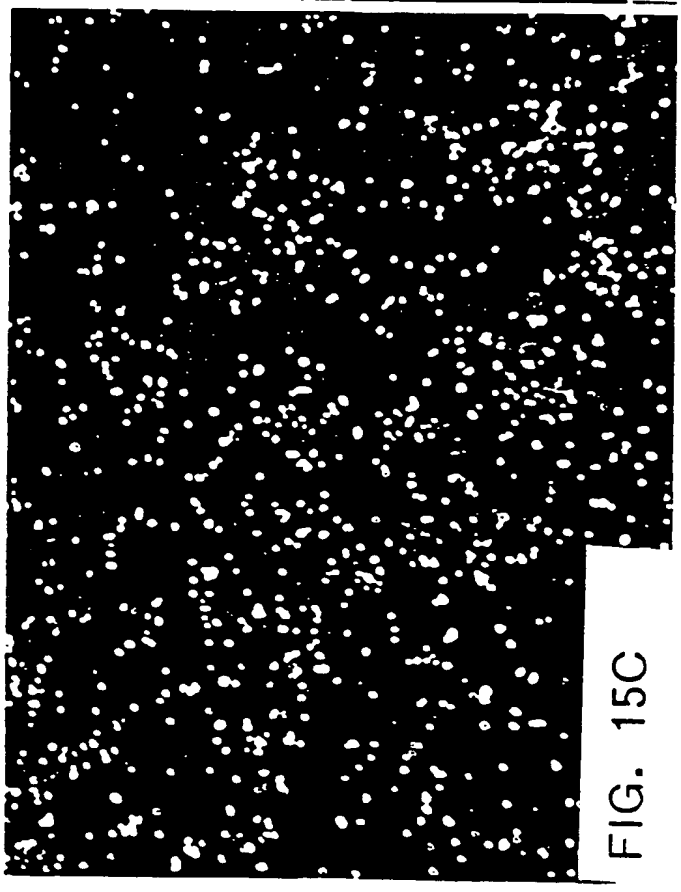


FIG. 15C

```
                                                    [  Transmembrane
obe2        FRNHPNITFFVYVSNFTWP--------IKIQIAFSQHSNFMDLVQFFVTFFSCFLSLLLVA
KIAA0534    FRSNPNITFYVYVSNFSWP--------IKIQIAFSQHNTIMDLVQFFVTFFSCFLSLLLVA
YC81_CAEEL  FGPDSNTTFFVRVYNFNTP--------VQIVVSFAQSPPIN-WVLFFVIFAACFIVLLVVA
MEGF8       LKSSRFYLLLLGVGDPSGPGANGSADSQGLLFFRQDQAHIDLFVFFSVFFSCFFLFLSLC
                                  Site
                                  ]
obe2        AVVWKIKQSCWASRRREQLLREMQQMASRPFASVETLPWNR----------------------
KIAA0534    AVVWKIKQTCWASRRREQLLRERQQMASRPFASVDVALEVGAEQTEFLRGPLEGAPKPIA
YC81_CAEEL  GLLWMIKVRIEAYRRNQRRIDEIEHMASRPFASTKMELSMLSQFSSAG-------------
MEGF8       VLLWKAKQALDQRQEQRRHLQEMTKMASRPFAKVTVCFPPDPTAPASAWKP-AGLPPP-A
                                  ********
```

FIG. 16A

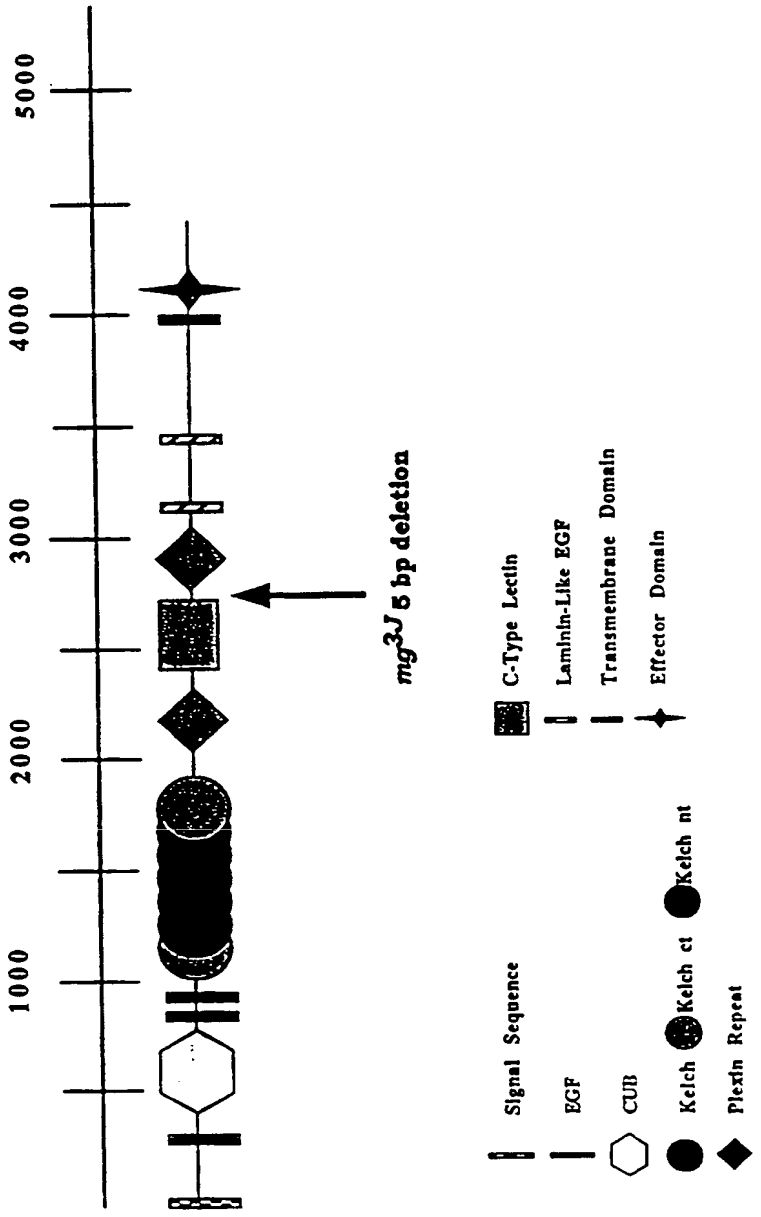


FIG. 16B

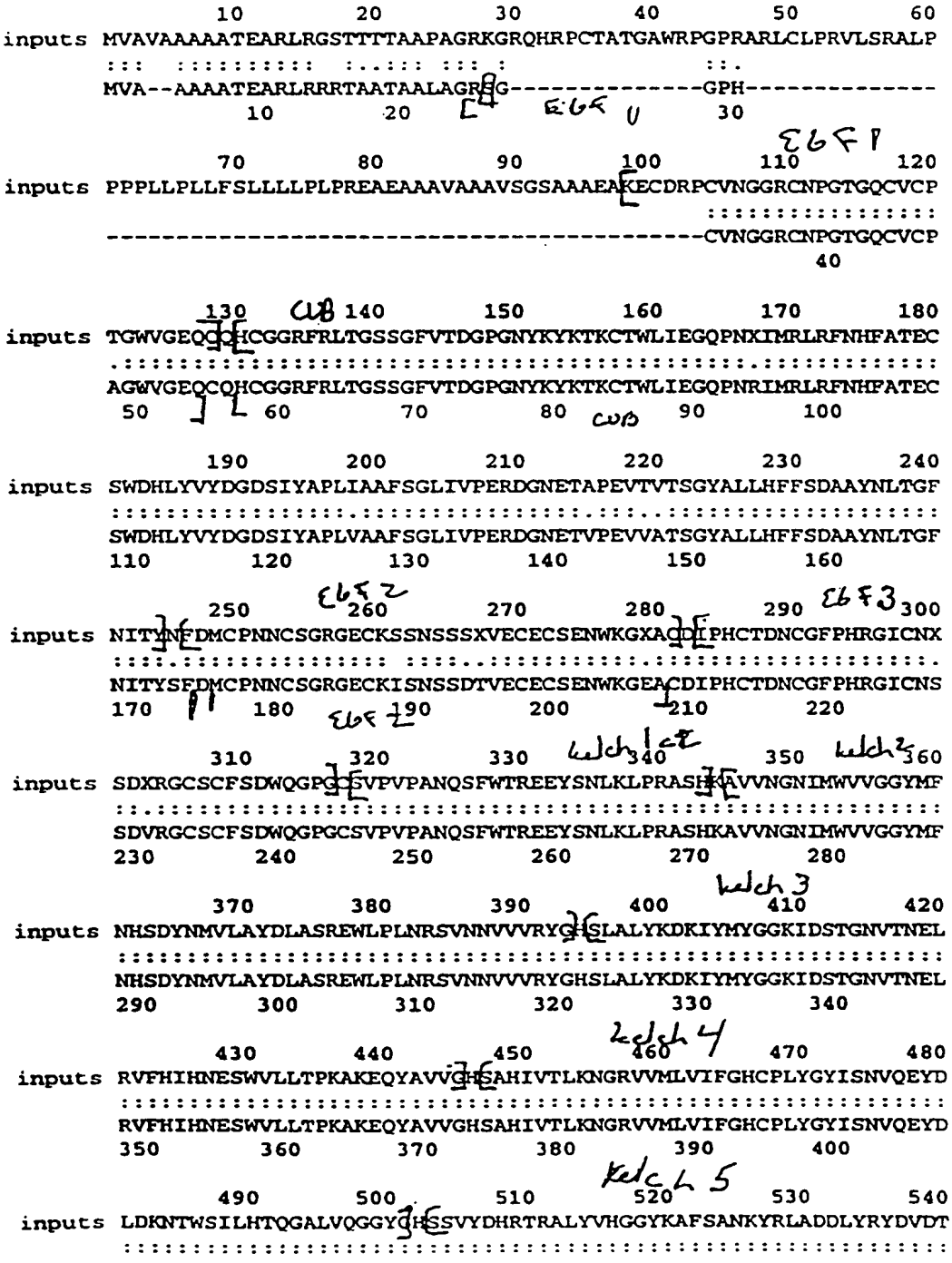


FIG. 17A

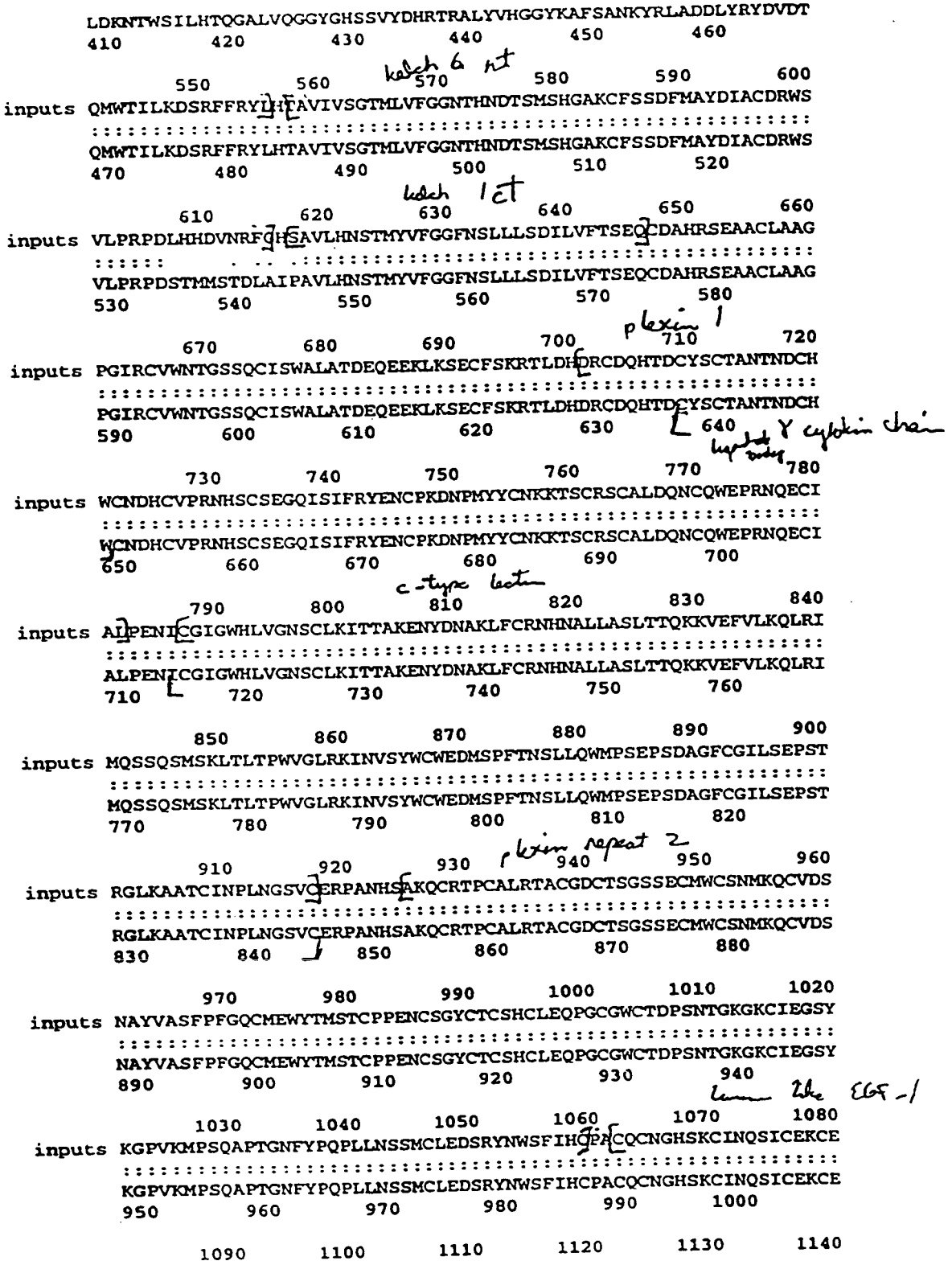

LDKNTWSILHTQGALVQGGYGHSSVYDHRTRALYVHGGYKAFSANKYRLADDLYRYDVDT
410 420 430 440 450 460
kelch 6 rpt
550 560 570 580 590 600
inputs QMWTILKDSRFFRYLHTAVIVSGTMLVFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWS
QMWTILKDSRFFRYLHTAVIVSGTMLVFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWS
470 480 490 500 510 520
kelch 1st
610 620 630 640 650 660
inputs VLPRPDLHHDVNRFGHSAVLHNSTMYVFGGFNSLLLSDILVFTSEQCDAHRSEAACLAAG
VLPRPDSTMMSTDLAIPAVLHNSTMYVFGGFNSLLLSDILVFTSEQCDAHRSEAACLAAG
530 540 550 560 570 580
plexin 1
670 680 690 700 710 720
inputs PGIRCVWNTGSSQCISWALATDEQEEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCH
PGIRCVWNTGSSQCISWALATDEQEEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCH
590 600 610 620 630 640
Y cytokine chain
730 740 750 760 770 780
inputs WCNDHCVPRNHSCSEGQISIFRYENCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECI
WCNDHCVPRNHSCSEGQISIFRYENCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECI
650 660 670 680 690 700
c-type lectin
790 800 810 820 830 840
inputs ALPENICGIGWHLVGNSCLKITTAKENYDNAKLFCRNHNALLASLTTQKKVEFVLKQLRI
ALPENICGIGWHLVGNSCLKITTAKENYDNAKLFCRNHNALLASLTTQKKVEFVLKQLRI
710 720 730 740 750 760
850 860 870 880 890 900
inputs MQSSQSMSKLTLTPWVGLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPST
MQSSQSMSKLTLTPWVGLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPST
770 780 790 800 810 820
plexin repeat 2
910 920 930 940 950 960
inputs RGLKAATCINPLNGSVCERPANHSAKQCRTPCALRTACGDCTSGSSECMWCSNMKQCVDS
RGLKAATCINPLNGSVCERPANHSAKQCRTPCALRTACGDCTSGSSECMWCSNMKQCVDS
830 840 850 860 870 880
970 980 990 1000 1010 1020
inputs NAYVASFPFGQCMEWYTMSTCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSY
NAYVASFPFGQCMEWYTMSTCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSY
890 900 910 920 930 940
like EGF-1
1030 1040 1050 1060 1070 1080
inputs KGPVKMPSQAPTGNFYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCE
KGPVKMPSQAPTGNFYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCE
950 960 970 980 990 1000
1090 1100 1110 1120 1130 1140

FIG. 17B

```
inputs NLTTGKHCETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCE
       :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
       NLTTGKHCETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCE
      1010      1020      1030      1040      1050      1060

            1150      1160      1170      1180      1190      1200
inputs VENRYQGNPLRGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASK
       :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
       VENRYQGNPLRGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASK
      1070      1080      1090      1100      1110      1120

            1210      1220      1230      1240      1250      1260
inputs NFNLNITWAASFSAGTQAGEEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFT
       :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
       NFNLNITWAASFSAGTQAGEEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFT
      1130      1140      1150      1160      1170      1180

            1270      1280      1290      1300      1310      1320
inputs WPIKIQIAFSQHSNFMILVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQLLREMQ
       :::::::.                                  :.        ::
       WPIKIQV-----------------------------------QT--------EQ------
      1190

            1330      1340      1350      1360      1370      1380
inputs QMASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRLPRGLGG
       ------------------------------------------------------------


            1390      1400      1410      1420
inputs IPPPGQSGLAVASALVDISQQMPIVYKEKSGAVRNRKQQPPAQPGTCIN
       ------------------------------------------------
```

FIG. 17C

ATGGTGGCGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAGGAGCGGCGGGCC
GCACCGACCCTGCACCGCGACAGGGGCCTGGAGGCCGGGACCGCGCGCCCGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTGCCCCCGC
CGCCGCTGCTGCCGCTGCTCTTTTCGCTGCTGCTGCTGCCGCTGCCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCG
GCCGCAGCCGAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTG
GGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACA
AAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCAT
TTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGT
CCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGTT
TTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGG
AAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTC
CTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGC
TCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTT
CTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATA
CAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCAT
GGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATG
CTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACAC
CCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTT
TCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTC
CGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAA
ATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGAT
TTGGCCATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCG
GAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTAT
CTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGC
ACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGC
CAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGA
CCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCAT
GTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAG
AAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCG
GAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTG
GATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCT
GCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTG
CAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCC
CCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAA
GGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTC
CAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGA
GCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAA
TGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGA
GTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCT
TTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATC
AATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAA
AACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTA
ATTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGT
TTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGAT

FIG. 18A(1)

GCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGA
CTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGC
ATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGG
AGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATCTGATGCTGGGGCCAGGGACTCTCCCACGCACGAGCTAGTG
AGTGGCACACCAGAGCCATCTGCAGGGAAGGGCGTGGCGGGGAAATGGCTGTGCGGTGCGGGACGGAAGACTGGAAACCCTCAAAGCATCTG
ACTCACCTGCATGATCACAAGCTTTCTTTGACGGTTTCTCCCATCCGTGTTCCAGCATCTAACCTTTTACTTTTGCATAGGAAATACTTGAT
TTAATTACAGGTCCAGGGATGAGCTGATGGTTGCTGGAGGAGGCCAGTGTAGAGCCAGTGAGAGAACTAGGAATGACACTCAGGTTCACTGT
GGAAAACTGTTCTTGGGACTGTCTCAACTGTGCAAAAAACAAAAGATGGAGTGTTTACAAGTAGACATTCGTCATCAGTTGTTCTTGAACAT
GGTCTTTTAAAAACTAGTCAGATGAATTAACTTGTTTTCATCTGAAGCCTGCTATCTTTTTTAAAAGATGTGCTATTTATTCTTGCACGATT
TAGGCAATTATCTCTCTTCCAGGGAGTACCTTTTTTTCTAGTTGAGAATTAATAATGGTCCATCTCTTTTGATCATATCAAGCTAGGATAGA
AGGGGGGCTATTTTAAATGTCAAGGTCAGCAGTGTTACTTTGAATGTAAACTGGTATAATAGGTAGTTTTCTATAGTAACTTGATTAATTTA
GTCTTAATCCATTTGAAACTCTCTCTTCCTTTCTCTCTGCCTGTCCCTCTCCTTCTCCATCTCACCCTCCCTCTCTCACACATACACACACA
AACACATACACACAACACTAAGTGCCTAGACTTTAAATAGATCTAGCAATTGGAAAGTTAGTAAGCCTAAGTTTTTACATAATTGCATTCCT
ACATTCTTGTAAAATTTAAATAGCTACCATTGGCAATCTGCTTTTTTTCTAAAATCTGATTTGCAGCCAGGAAAGAATTTTCTCACCCAAGG
AACATTTGATCTAGCAGCAGGGATGAGAGGAAAGCAGAAATGAATGAACTGTGAAAGCTCCTGTTTTTATTATCAAAAAGGACACTGTCAAG
AAGGCGCCCCCTGCCCCCACCCCCGTGTCACCCTAGGCCTGATAAGCGATCAGAGGAAAGGACTCATTCATGTCACGCTTCCTTGAGCAGAA
AAGAGCACTGAGAGCACTTGGGACCCCTGGATCAGAGAGCATCTGTGTGTCCTGCAGCCTCCTCTGAACTTGTGGTTCATTCTCAGGCTGGG
GTGGACTCAGATGCCAGGAAAGGGACAGCCTCCCATTGTCAGGCAGAAGCTGCCCAAAGCCTGGAGAAGGACTTGTTTGCCCTCTTTCCCCC
AGGAGGGGCTCGACCCACCCACCCTCCCTCTCAGACCAAGGTGGTGGCTGTGAGGAGGGCAGCAAATGCTGACAAGGATGAAAAGCACATGG
AAAAAAATGGACGAGGAGGGAAAACTCTGCCAAATGGAAAATGACCAAATTTAAGAGGGTGGGACAGTCCCCTGCTCCTCTCCCAGAGGGCA
CTGCTTGGAAATTGTGTTTTCCCCATTTATGGTGCTCTGTATTCTGGCATTATGCAGCAGCCTCCCAGAAGCTCTCTTCTGCTTCAAAACCT
GGGATCTCTGGCATTACCCTATTGGGATGGACCGCTGGACAGCAATGCTCGAGTTTGTGAATTTGGAGAGATACTCAAAAGAGCTAAAACTG
CAGCATTTTACCTTTAAATGCAGTGCCTAGAGAGAGAGTATTGTCTCTTCCCCAACACTAACCCCACTCCCATGAAGAATTGCCTGGAAAGA
TGTTTTCAAGGAATTTGAACCATAAAACACTATCTGATGCACAGAACACCTCTACTTTGAGACTCACCTCTCATAAAGCTTCTTTTTCACAT
TACTGTTAAAGACCAGACGTTCTAGAAAAGACCCCTCCTCTCATGAGCTCCCCCATCCCTGCTACAGAACACAGCACCCATGGCGCCTGCAG
TGGACTGGCCCCTTAATTCCCACAGGCCCCCCCAGCAAGGCCAAAGGGAGGCCCCTGGGTATTGTCCTCCTACAAGGAAGATCCTCTTTGTT
TGTTCAAAGGACCAGTTTTCCTAGGCCAAAGAAGTCTCTTCCCCATGTTAGTCCTATGCCTTGAAATATCATGCACCATGACCCACAGCCAT
CTGGTTATGTCTTATTTTTTTCCTAAAAGATAATGTTTATTTTTAAAAAGGAAGGAAGAAGCAAGTGAAGTTTCATTCTGCTCCAGCGGTGG
GGAAGCCGCTGAATCCACCTGCTTCTCCTTTGCAACCGACAGCAAACAGCTTTCTCCGGCCTCAGGGCAGAAAAAGGGAATGGCAGGGAGTA
AGAGGCGCTGGGCTCGGAGCCTGTTTCCAAGAAGGAATTGGTTGTCATCTGGCAGTGTTGCGCGTCACAAGAGAGCCTGTATATAAATTAAA
ATAGTCAAGACAACACTGACCTTGCACTTGTACATAACTATACAGTAGTGTCCAGAATGTTCAGACATTCGGAGTGTACATAAAACAGAAAA
AATCTTCATGTATTTTTATTAAATATAACAATGTCTGAGTTTCACCTAAGATGTTTTTGTGCCATATGCTGGATATCCAGGTTCTCGCCAGG
CCCCGATACATGAATAACAAACCCAAGAAACGCATCCCCATTGTGTGATGTGTTCAGATGCATCTGGCACCAATTAGGTATTTCTTAAAACA
GGACTCATCTGTCAGAGTGCACATGAAAAATCAGGCAGGGAATCGAAACGACAGCGCTGGAGGAGACTCAGGAAGCAGAGGCGTCCCTGCCG
CTGCCCTTGGCCCTGCAAGCACATCATGACCCTTTCTGGCAGCCTCTTGGTGCTCTGGGTAGTGAGGGATGACCAGTCTTGTCCTGAGAAAT
GTTTCTCTTAGTCTTTAAGTTCAAAGACTAACCTGTAGCAATCAGACTTTCCAAAAGGGGGTTCTCCATTTTTTGTAGTTTTGTCTAAATTT
TTAATGACCATTTCCTGGAATCAGTTTATTATACTGAAAACTGGGGGTGGGAGTAGGGAGCTAGTTTGTTGATAAATAGTTCCCATTTCCCC
GTGGAGAATTTGACATACCCTGGACTCCTGTGTGCCTCCTGCCATCCCTGCACACAGCCTGGGGAGAAGCCTGTGCCTCCCCGTGTGGAGAG
AAGGCAACCCCAGATCCCCTGAGCTAACCCGGAGGAAAGGCAGTCCTGGACAGAAGACTGTCAGCAGAAGGAAAGTACTGGACTACCCGTGG
GTAAGTCCTGCCATTCAAGACTGGAGACACCTGGGAAATAAAAAGAGCAGGGCACTGCTGGTGGGAAGAGGCATTTTACCTTCCAGTGCAAA
TCCTGCTCCTTTGATTTAATGGGGTGTACTGGGGCCAGGGGCTGATTCACTTCCTTGGGAGATGGTGGTGTTTTCATGAACATCTTTGATCC
TTCCATTTCATTTATTCATCCATCCATTCAACAAGTATTTGCTAAACACTAACTTAAGCTAATGCTAGGGTAGTGACTGAGATGTAAAAATA
GATTTTAGAATTAAAACAAAATCCAAGTCCTCACACCCCTGTCATCCCAGGAGATCTTTCCTTGTGGTGGTTTCTGTGAGAATTGGCCATCC

FIG. 18A(2)

TGAGGACACAGCCAGGACGGCAGAGGCCTCCTGGCCTCAGGGCATGCCCTGCCTACCTTCTGAAATGTTTACCCCATTGACCAAACTTGGCT
CCAGCCATTGCGGTGGTTTCTAGATAGCCAGGCCCACCAAGAGATATTGCCCCTTGATGAGAGTCAAACACCCTGCCTACAAGGAGATGTTT
TGAAATGGAGAGGAAAATTGGCACCTCATCTTTTAAAGGCAGTAATGGAATTGATTTTCAGTAACTGAATTTGTGCACAAAACATTCTAAAC
ACTAGTGAAGCCTGTTTCGTTGAACTAATTCTGGCTCTGGAAATGTTTTTGTTTTATAGTTATTTACGATTTCGTTTGTTTGGATTCAAGCT
TAGTTTGTTAATATGTATAATTTAGCATCTATTACACTCATGTAAATATGGAGTAAGTATTGTAAACTATTTCATTGCGGGGATTGTGGGTG
TTATACATACATTTAGGACTGCAATTTTTTGGTATTTTTTGTATTGTAAAATAACAGCTAATTTAAGCAGGAACAAGAGAACTAAGGGAGGT
CTGTGCATTTTAAACACAAATGTGAAGAACTTGTATATAAACAAAAGTAAATACTATAATACAAACTTCCTTCTGAAATAAAAGTAGATCTG
GTAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAA

FIG. 18A(3)

MVAVAAAAATEARLRRRTAATAALAGRSGGPHRPCTATGAWRPGPRARLCLPRVLSRALPPPPLLPLLFSLLLLPLPREAEAAAVAAAVSGS
AAAEAKECDRPCVNGGRCNPTTGQCVCPAGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTWLIEGQPNRIMRLRFNHFATECSWDH
LYVYDGDSIYAPLVAAFSGLTVPERDGNETVPEVVATSGYALLHFFSDAAYNLTGFNITYSFDMCPNNCSGRGECKISNSSDTVECECSENW
KGEACDIPHCTDNCGFPHRGTCNSSDVRGCSCFSDWQGPGCSVPVPANQSFWTREEYSNLKLPRASHKAVVNGNIMWVVGGYMFNHSDYNMV
LAYDLASREWLPLNRSVNNVMVRYGHSLALYKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKEQYAVVGHSAHIVTLKNGRVVM
LVIFGHCPLYGYISNVQEYDLDKNTWSILHTQGALVQGGYGHSSVYDHRTRALYVHGGYKAFSANKYRLADDLYRYDVDTQMWTILKDSRFF
RYLHTAVIVSGTMLVFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPDLHHDVNRFGHSAVLHNSTMYVFGGFNSLLLSDILVFTS
EQCDAHRSEAACLAAGPGIPTVWNTGSSQCISWALATDEQEEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPRNHSCSEG
QISIFRYENCPKDNPMYYCNPKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLVGNSCLKITTAKENYDNAKLFCRNHNALLASLTTQ
KKVEFVLKQLRIMQSSQSMSVLTLTPWVGLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNGSVCERP
ANHSAKQCRTPCALRTACGLCTSGSSECMWCSNMKQCVDSNAYVASFPFGQCMEWYTMSTCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGK
GKCIEGSYKGPVKMPSQAPTGNFYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCENLTTGKHCETCISGFYGDPTNGGK
CQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQGNPLRGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFI
NASKNFNLNITWAASFSAGTQAGEEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPIKIQIAFSQHSNFMDLVQFFVTFFSC
FLSLLLVAAVVWKIKQSCWASRRREQLLREMQQMASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRLPRGLGG
IPPPGQSGLAVASALVDISQQMPIVYKEKSGAVRNRKQQPPAQPGTCICWGQGLSHARASEWHTRAICREGRGGEMAVRCGTEDWKPSKHLT
HLHDHKLSLTVSPIRVPASNLLLLLHRKYLILQVQGADGCWRRPVSQENEHSGSLWKTVLGTVSTVQKTKDGVFTSRHSSSVVLEHGLLKTSQ
MNLVFISLLSFLKDVLFILAPFRQLSLFQGVPFFLVENWSISFDHIKLGKGGYFKCQGQQCYFECKLVVVFYSNLINLVLIHLKLSLPFSLP
VPLLLHLTLPLSHIHTQTHTHNTKCLDFKIQLESAVFTLHSYILVKFKLPLAICFFSKIFAARKEFSHPRNISSSRDERKAEMNELKLLFLL
SKRTLSRRRPLPPPPCHPRPDKRSEERTHSCHASLSRKEHEHLGPLDQRASVCPAASSELVVHSQAGVDSDARKGTASHCQAEAAQSLEKDL
FALFPPGGARPTHPPSQTKVVAVRRAANADKDEKHMEKNGRGGKTLPNGKPNLRGWDSPLLLSQRALLGNCVFPIYGALYSGIMQQPPRSSL
LLQNLGSLALPYWDGPLDSNARVCEFGEILKRAKTAAFYLMQCLEREYCLFPNTNPTPMKNCLERCFQGITIKHYLMHRTPLLDSPLIKLLF
HITVKDQTFKRPLLSAPPSLLQNTAPMAPAVDWPLNSHRPPQQGQREAPGYCPPTRKILFVCSKDQFSAKEVSSPCSYALKYHAPPTAIWLC
LIFFLKDNVYFKGRKKQVKFHSAPAVGKPLNPPASPLQPTANSFLRPQGRKREWQGVRGAGLGACFQEGIGCHLAVLRVTREPVYKLKSRQH
PCTCTLYSSVQNVQTFGVYIKQKKSSCIFIKYNNVVSPKMFLCHMLDIQVLARPRYMNNKPKKRIPIVCVQMHLAPIRYFLKQDSSVRVHMK
NQAGNRNDSAGGDSGSRGVPAAALGPASTSPFLAASWCSGGMTSLVLRNVSLSLVQRLTCSNQTFQKGVLHFLFCLNFPFPGISLLYKLGVG
VGSFVDKFPFPRGEFDIPWTPVCLLPSLHTAWGEACASPCGEKATPDPLSPGGKAVLDRRLSAEGKYWTTRGVLPFKTGDTWEIKRAGHCWW
EEAFYLPVQILLLFNGVYWGQGLIHFLGRWWCFHEHLSFHFIYSSIHSTSICTLTANARVVTEMKILELKQNPSPHTPVIPGDLSLWWFLEL
AILRTQPGRQRPPGLRACPAYLLKCLPHPNLAPAIAVVSRPGPPRDIAPESNTLPTRRCFEMERKIGTSSFKGSNGIDFQLNLCTKHSKHSL
FRTNSGSGNVFVLLFTISFVWIQAFVNMYNLASITLMIWSKYCKLFHCGDCGCYTYIDCNFLVFFVLNNSFKQEQENGRSVHFKHKCEELVY
KQKILYKLPSEIKVDLVKKKKEKKKK

FIG. 18B

ATGGTGGCGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCCGCGACGGCAGCGCTCGCGGGCAGGAGCGGCGGGCCGCACCGACCCTGCACC
GCGACAGGGGCCTGGAGGCCGGGACCGCGCGCCCCGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTGCCCCCGCCGCCGCTGCTGCCGCTGCTCTTTTCGCTGCTG
CTGCTGCCGCTGCCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGCTCGGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGT
CGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTG
ACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGT
AGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCT
GAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAAC
TGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGAC
AACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAAC
CAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTC
AACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTG
GCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTG
TTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCT
CTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGT
GTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACC
CAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACA
TCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAAC
AGATTTGGCCATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGAT
GCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAA
GAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCAC
TGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAAC
AAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCAT
TTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAG
AAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCC
TACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACT
CGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA
GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGT
ATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGC
AATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGC
ATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAG
AACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTG
TGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGA
ACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAA
AACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTT
GTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACC
TGGCCCATCAAAATTCAGGTGCAAACTGAACAATGAGGACGCATGGACACAGGAAGGGGAACATCACACACCAGGGCCTGTTGTGGGGTGGGGGGAAGGGGAAGGGAT
AGCATTAGGGGATATACCTAATGTTAAATGACGAGTTAATGGGTGCAGCACACCAACATGGCATATGTATACATATGTAACAAACCTGCATGTTGTGCACATGTACCC
TAAAACTTAAAGTATAATTAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIG. 19A

MVAVAAAAATEARLRRRTAATAALAGRSGGPHRPCTATGAWRPGPRARLCLPRVLSRALPPPPLLPLLFSLLLLPLPREAEAAAVAAAVSGSAAAEAKECDRPCVNGG
RCNPGTGQCVCPAGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTWLIEGQPNRIMRLRFNHFATECSWDHLYVYDGDSIYAPLVAAFSGLIVPERDGNETVP
EVVATSGYALLHFFSDAAYNLTGFNITYSFDMCPNNCSGRGECKISNSSDTVECECSENWKGEACDIPHCTDNCGFPHRGICNSSDVRGCSCFSDWQGPGCSVPVPAN
QSFWTREEYSNLKLPRASHKAVVNGNIMWVVGGYMFNHSDYNMVLAYDLASREWLPLNRSVNNVVVRYGHSLALYKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVL
LTPKAKEQYAVVGHSAHIVTLKNGRVVMLVIFGHCPLYGYISNVQEYDLDKNTWSILHTQGALVQGGYGHSSVYDHRTRALYVHGGYKAFSANKYRLADDLYRYDVDT
QMWTILKDSRFFRYLHTAVIVSGTMLVFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPDLHHDVNRFGHSAVLHNSTMYVFGGFNSLLLSDILVFTSEQCD
AHRSEAACLAAGPGIRCVWNTGSSQCISWALATDEQEEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPRNHSCSEGQISIFRYENCPKDNPMYYCN
KKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLVGNSCLKITTAKENYDNAKLFCRNHNALLASLTTQKKVEFVLKQLRIMQSSQSMSKLTLTPWVGLRKINVS
YWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNGSVCERPANHSAKQCRTPCALRTACGDCTSGSSECMWCSNMKQCVDSNAYVASFPFGQC
MEWYTMSTCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQAPTGNFYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCE
NLTTGKHCETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQGNPLRGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQ
NRDLDMFINASKNFNLNITWAASFSAGTQAGEEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPIKIQVQTEQGRMDTGRGTSHTRACCGVGGRGRDS
IRGYTCMTSWVQHTNMAYVYICNKPACCAHVPNLKYNKKKKKKKKKKKKKKKKKKK

FIG. 19B

ATGGTGGCGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAGGAGCGGCGGGCCGCACCGACCCTGCACC
GCGACAGGGGCCTGGAGGCCGGGACCGCGCGCCCGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTGCCCCCGCCGCCGCTGCTGCCGCTGCTCTTTTCGCTGCTG
CTGCTGCCGCTGCCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGT
CGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTG
ACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGT
AGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCT
GAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAAC
TGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGAC
AACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAAC
CAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTC
AACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTG
GCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTG
TTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCT
CTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGT
GTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACC
CAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACA
TCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAAC
AGATTTGGCCATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGAT
GCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAA
GAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCAC
TGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAAC
AAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGGTAGGCCTTGCAGGGTCATCTTGGTG
TGTGTGGGTCCATTACTTCAGCCTGCTTCCCCCAACACTGTGCAGCCTAAGTTGAACCTAGCAGAGGGGAAGAGCTAATTCTGTCCATTCATCCCCCACACGAGTATT
ATGGGCTTTTTTGTTTTTAACTAAAATACAGTTCTTAAGTATTTGTTCCTACTGTCCTTTGAAATAAAGTGAAACATCCTTTGCTGCTCTGTAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIG. 20A

MVAVAAAAATEARLRRRTAATAALAGRSGGPHRPCTATGAWRPGPRARLCLPRVLSRALPPPPLLPLLFSLLLLPLPREAEAAAVAAAVSGSAAAEAKECDRPCVNGG
RCNPGTGQCVCPAGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTWLIEGQPNRIMRLRFNHFATECSWDHLYVYDGDSIYAPLVAAFSGLIVPERDGNETVP
EVVATSGYALLHFFSDAAYNLTGFNITYSFDMCPNNCSGRGECKISNSSDTVECECSENWKGEACDIPHCTDNCGFPHRGICNSSDVRGCSCFSDWQGPGCSVPVPAN
QSFWTREEYSNLKLPRASHKAVVNGNIMWVVGGYMFNHSDYNMVLAYDLASREWLPLNRSVNNVVVRYGHSLALYKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVL
LTPKAKEQYAVVGHSAHIVTLKNGRVVMLVIFGHCPLYGYISNVQEYDLDKNTWSILHTQGALVQGGYGHSSVYDHRTRALYVHGGYKAFSANKYRLADDLYRYDVDT
QMWTILKDSRFFRYLHTAVIVSGTMLVFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPDLHHDVNRFGHSAVLHNSTMYVFGGFNSLLLSDILVFTSEQCD
AHRSEAACLAAGPGIRCVWNTGSSQCISWALATDEQEEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPRNHSCSEGQISIFRYENCPKDNPMYYCN
KKTSCRSCALDQNCQWEPRNQECIALPGRPCRVILVCVGPLLQPASPNTVQPKLNLAEGKSFCPFIPHTSIMGFFVFNNTVLKYLFLLSFEIKNILCCSVKKKKKKKK
KKKKKKKK

FIG. 20B

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/12, C07K 14/705, A61K 38/17, G01N 33/68, C07K 16/28

A3

(11) International Publication Number: WO 00/05373

(43) International Publication Date: 3 February 2000 (03.02.00)

(21) International Application Number: PCT/US99/16484

(22) International Filing Date: 21 July 1999 (21.07.99)

(30) Priority Data:

60/093,630	21 July 1998 (21.07.98)	US
60/104,978	20 October 1998 (20.10.98)	US
09/245,041	5 February 1999 (05.02.99)	US

(71) Applicant: MILLENIUM PHARMACEUTICALS, INC. [US/US]; 640 Memorial Drive, Cambridge, MA 02139 (US).

(72) Inventors: MOORE, Karen; 34 Chandler Street, Maynard, MA 01754 (US). NAGLE, Deborah, L.; 370 Arlington Street, Watertown, MA 02172 (US).

(74) Agents: CORUZZI, Laura, A. et al.; Pennie & Edmonds LLP, 1155 Avenue of the Americas, New York, NY 10036 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(88) Date of publication of the international search report:
29 June 2000 (29.06.00)

(54) Title: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY

(57) Abstract

The present invention relates to mammalian mahogany genes, including the human mahogany gene, which are novel genes involved in the control of mammalian body weight. The invention encompasses nucleotide sequences of the mahogany gene, host cell expression systems of the mahogany gene, and hosts which have been transformed by these expression systems, including transgenic animals. The invention also encompasses novel mahogany gene products, including mahogany proteins, polypeptides and peptides containing amino acid sequences mahogany proteins, fusion proteins of mahogany proteins polypeptides and peptides, and antibodies directed against such mahogany gene products. The present invention also relates to methods and compositions for the diagnosis and treatment of mammalian body weight disorders, including obesity, cachexia, and anorexia, and for the identification of subjects susceptible to such disorders. Further, the invention relates to methods of using the mahogany gene and gene products of the invention for the identification of compounds which modulate the expression of the mahogany gene and/or the activity of the mahogany gene product. Such compounds can be useful as therapeutic agents in the treatment of mammalian body weight disorders, including obesity, cachexia, and anorexia.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

INTERNATIONAL SEARCH REPORT

Interr ıal Application No
PCT/US 99/16484

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/705 A61K38/17 G01N33/68 C07K16/28

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K C12N A61K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	NAGASE ET AL.: "Prediction of the coding sequences of unidentified human genes IX: the complete sequences of 100 new clones from brain which can code for large proteins in vitro" DNA RESEARCH, vol. 5, 1998, pages 31-39, XP000884356 table 2 ---	1,2,5,11
X	DATABASE GENBAN 'Online! Accession no. AB11120, 10 April 1998 (1998-04-10) NAGASE ET AL.: "Prediction of the coding sequences of unidentified human genes" XP002135391 abstract --- -/--	1,2,5,11

[X] Further documents are listed in the continuation of box C. [ ] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
11 April 2000	27/04/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Skelly, J

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Interr ıal Application No
PCT/US 99/16484

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
P,X	NAGLE ET AL.: "The mahogany protein is a receptor involved in suppression of obesity" NATURE, vol. 398, 11 March 1999 (1999-03-11), pages 148-152, XP002135389 the whole document	1-28
P,X	GUNN ET AL.: "The mouse mahogany locus encodes a transmembrane form of attractin" NATURE, vol. 398, 11 March 1999 (1999-03-11), pages 152-156, XP002135390 the whole document	1-28
X	DUKE-COHAN ET AL.: "A novel form of dipeptidylpeptidase IV found in human serum" J. BIOL. CHEM., vol. 270, no. 23, 9 June 1995 (1995-06-09), pages 14107-14114, XP000579864 page 14109 -page 14111	16-18

4

INTERNATIONAL SEARCH REPORT

In. .iational application No.
PCT/US 99/ 16484

Box I Observations wh re certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for th following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although claims 26-28 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additi nal search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.